

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:14:27 ; Search time 20.83 Seconds

(without alignments)
1437.746 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598
Sequence: 1 MAASMGVLQEREFKKIWT.....KLVRTTEGQWGLFTAKK 494

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0601:*

1:	/SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2204	84.8	473	22	AA80627
2	206.5	7.9	285	21	AA770038
3	204.5	7.9	317	20	AAW95016
4	198	7.6	283	20	AA339308
5	193.5	7.4	279	21	AA170041
6	189.5	7.3	275	21	AA32508
7	189.5	7.3	283	21	AA32507
8	188	7.2	293	21	AA321652
9	188	7.2	296	21	AA321526
10	188	7.2	336	21	AA321651
11	188	7.2	339	21	AA321525

12	188	7.2	361	21	AA321524	Arabidopsis thalia
13	188	7.2	366	21	AA321520	Arabidopsis thalia
14	187	7.2	293	21	AA321522	Arabidopsis thalia
15	187	7.2	293	21	AA321525	Arabidopsis thalia
16	187	7.2	336	21	AA321521	Arabidopsis thalia
17	187	7.2	336	21	AA321524	Arabidopsis thalia
18	187	7.2	358	21	AA321520	Arabidopsis thalia
19	187	7.2	365	21	AA321523	Arabidopsis thalia
20	185	7.1	301	22	AA321523	Synchocystis thalia
21	185	7.1	318	21	AA321524	Synchocystis thalia
22	185	7.1	318	22	AA321523	Synchocystis thalia
23	185	7.1	318	22	AA321523	Synchocystis thalia
24	182	7.0	308	21	AA321523	Arabidopsis thalia
25	182	7.0	344	21	AA321523	Arabidopsis thalia
26	182	7.0	348	21	AA321523	Arabidopsis thalia
27	179	6.9	280	18	AA321523	S-adenosyl methion
28	178	6.9	348	20	AA321523	Arabidopsis thalia
29	178	6.9	348	21	AA321523	Arabidopsis thalia
30	177	6.8	308	21	AA321523	Arabidopsis thalia
31	177	6.8	348	21	AA321523	Arabidopsis thalia
32	173.5	6.7	266	21	AA321523	Arabidopsis thalia
33	172	6.6	115	21	AA321523	SAM dependent meth
34	170	6.5	352	21	AA321523	Corn gamma-tocophe
35	170	6.5	383	20	AA321523	Yeast SMT enzyme.
36	169.5	6.5	114	21	AA321523	SAM dependent meth
37	168.5	6.5	275	20	AA321523	Arabidopsis thalia
38	168	6.5	275	20	AA321523	Spnf protein invol
39	165	6.4	349	21	AA321523	Soybean gamma-toco
40	164.5	6.3	350	21	AA321523	Soybean gamma-toco
41	162.5	6.3	344	20	AA321523	Corn SMT enzyme.
42	162.5	6.3	366	21	AA321523	Soybean gamma-toco
43	162.5	6.3	581	21	AA321523	Arabidopsis thalia
44	160	6.2	361	20	AA321523	Arabidopsis SMT en
45	159.5	6.1	357	20	AA321523	Maize cyclopropane

ALIGNMENTS

RESULT	1
AA80627	standard; Protein: 473 AA.
ID	AA80627
AC	AA80627
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Environmental stress tolerant protein SEQ ID 40.
XX	
KW	Environmental stress resistance; salt; heat; desert; transgenic plant.
XX	
OS	Suada japonica.
XX	
PN	WO200106006-A1.
XX	
PD	25-JAN-2001.
XX	
PE	19-JUL-2000; 2000WO-JP04862.
XX	
PR	19-JUL-1999; 99JP-0235910.
XX	
PR	24-MAR-2000; 2000JP-0085377.
XX	
PA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX	
PI	Yamada A, Ozeki Y, Saito T;
XX	
DR	WPI; 2001-147355/15.
XX	
DR	N-PSDB; AAF74206.
XX	
PT	Screening method to obtain DNA encoding environmental stress resistance
XX	
PT	factor, useful for producing transgenic plants resistant to
XX	
XX	environmental stress

PS Claim 64; Page 125-127; 167pp; Japanese.
 XX
 CC Polynucleotide sequences AAF74187 - AAF74218 encode proteins
 CC AAB80608 - AAB80639, which impart environmental stress resistance. The
 CC invention relates to a method for identifying DNA encoding proteins
 CC imparting environmental stress resistance. The method comprises inserting
 CC cDNA from a library originating in a salt-resistant organism into a host
 CC cell, culturing the transformants under conditions in which the
 CC untransformed host does not grow well, and selecting for viable clones.
 CC The method is useful for obtaining DNA encoding environmental stress
 CC resistance factors. The DNA encoding proteins conferring environmental
 CC stress resistance, can be used in the production of plants resistant to
 CC environmental stress, which can be cultivated in unfavourable
 CC environments such as deserts, salt damaged ground, cold regions and the
 CC oceans. They can be used for increasing the area of land covered by green
 CC plants, and desert greening and afforestation. In order to counter the
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR
 CC primers AAF74219 and AAF74220 are used in an example illustrating the
 CC method of the invention.
 CC
 XX
 SO Sequence 473 AA;
 Query Match 84.8%; Score 2204; DB 22; Length 473;
 Best Local Similarity 86.4%; Pred. No. 3.9e-198;
 Matches 408; Conservative 34; Mismatches 30; Indels 0; Gaps 0;
 22 HSDVLTVEAMWDSQASDLKVERPEVLSMPREKSVLELNGIRFGELAKRSQV 81
 1 htvdlleamwdsqasdlkverpelismprlekswlelngirfgelaekrsqv 60
 82 IALDFIESVIRKKNESINGHYKKNVFCADYVSPSLNISPNVSDIFFNMLMYLSDDEVE 141
 61 ialdfiesalrknevinghyknvfcadvtspslsfpnsldvifsnwlmlylsdeve 120
 142 RLYERMLKMLKPGCYTRFRSCFQSGDHKRSNPTHYRPREFTTKIFKCHMODDSGNS 201
 121 rlyermlkmlkpgcytrfrscfqsghkrsnpthyrprefttkifkchmoqdsqns 180
 202 YELSLICGICGAVKSKKONONISMLMOKVYDSEDDKFOREFLSDSOYKFNSTIRYRVF 261
 181 yelslicgicgavvrnkknqnlswlwgkvdskdkqgflfildtsykcnsllryrvf 240
 262 GPGVSTGLETTEKFEVSKLDLKGQKVLVDCGIGGDFYMAENYVEVVGIDLSINMI 321
 241 gpgvstgylettekfevskldlkgqkvlvdcgigggdfymaenfvvvgidlsinmi 300
 322 SPALERSIGLKCAVEFEVADCTKKDIPENSFVDVYSRDTILIHODKPALFRSFHKMLKPG 381
 301 sfalersiglkcavefevadctkknypdsvfysrdtilihqdkpalfrsfhkmlkpg 360
 382 GKVLISDYCKSAGTSPSAFAYIRORGYD-LHDVKAQKMLKDAGEVVAENRTDQFIQ 441
 361 gkvlisdycskaagtpsaafayirorgydlhdvkaqkmlkdagevvaenrtdqfiq 420
 442 LQKELDALOEKDDFIDFSEEDYNDIVDGWAKALVTRTEGQWGLFIKK 493
 421 lrkeldelekekdfidfsfeedyndivdgwacklrrtagewglfvakk 472
 Db
 RESULT 2
 ID AAY70038 standard; protein; 285 AA.
 XX
 AC AAY70038;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE A. halophila sarcosine-dimethylglycine methyltransferase.
 KW Sarcosine-dimethylglycine methyltransferase; SDMT; betaine operon;
 KM dimethylglycine content; abiotic stress; tolerance; salt; freezing;
 KW cold; drought stress; pathogenesis-related protein; animal feeding;

KW pH tolerance; environmental stress; inclusion body formation.
 XX
 OS Actinopolyspora halophila.
 XX
 PN WO200011142-A2.
 XX
 PD 02-MAR-2000.
 XX
 PF 18-AUG-1999; 99WO-EP06037.
 XX
 PR 20-AUG-1998; 98US-0137434.
 XX
 PA (CULT-) CULTOR CORP.
 PI Reinikkaenen T, Nyssöelae A, Keroyno J;
 XX
 DR WPI; 2000-224686/19.
 DR N-PSDB; AA250960.
 XX
 PT New methyltransferases, useful for e.g. producing transgenic plants
 PT with increased stress tolerance, pathogen resistance or nutritional
 PT value as animal feed or for improving the viability of microorganisms
 PT in the food industry
 PS
 PS Claim 19; Page 150-152; 176pp; English.
 CC
 CC The present sequence is Actinopolyspora halophila sarcosine
 CC -dimethylglycine methyltransferase (SDMT) catalysing conversion of
 CC sarcosine to dimethylglycine. Expression vectors comprising the coding
 CC region from betaine operon can be used to increase intracellular betaine
 CC and dimethylglycine content. Betaine is accumulated in the cells as a
 CC result of abiotic stress, and the polynucleotide can be used to increase
 CC salt, freezing or cold tolerance, increase resistance to drought stress,
 CC pathogens or induce pathogenesis-related proteins in plants. Transgenic
 CC organisms can be used as an animal feed ingredient. The polynucleotide
 CC can also be used to enhance pH tolerance and improve viability of
 CC organisms when subjected to environmental stress. This can decrease
 CC inclusion body formation when used in conjunction with polynucleotides
 CC encoding a heterologous protein.
 CC
 XX
 SO Sequence 285 AA;
 Query Match 7.9%; Score 206.5; DB 21; Length 285;
 Best Local Similarity 26.3%; Pred. No. 4.5e-11;
 Matches 61; Conservative 44; Mismatches 120; Indels 7; Gaps 4;
 266 VSTGGLTTEKFEVSKLDLKGQKVLVDCGIGGDFYMAENYVEVVGIDLS--INMISF 323
 55 latsertvgmagkvdvspetrilldigayggaarylartgchvchlnlsevengrnr 114
 324 ALERSIGLKCAVEFEVADCTKKD--YFENSFVDVYSRDTILIHODKPALFRSFHKMLKPG 381
 115 eltraeglehl--evdgsfedipygdnafdvwsqdsflhsgrsrmveevrtvlpk 172
 382 GKVLISDYCKSAGTSPSAFAYIRORGYD-LHDVKAQKMLKDAGEVVAENRTDQFIQ 440
 173 gsvlftdpmasdaakneipldrilnldslgspgfyfkrkeltrtlglqnlsfedlseylpv 232
 441 VLOKELDALOEKDDFIDFSEEDYNDIVDGWAKALVTRTEGQWGLFIKK 492
 233 hgyrlevlesreneleagfigeeyrahmkcglrnwvgagsgslawgljnar 284
 Db
 RESULT 3
 ID AAW95016 standard; protein; 317 AA.
 XX
 AC AAW95016;
 XX
 DT 21-MAY-1999 (first entry)
 XX
 DE Synechocystis gamma-tocopherol methyltransferase (gamma-TMT).

```

XX  Gamma-tocopherol methyltransferase; gamma-TMT; SLR0089; 165H5T7; meat;
KW  transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant.
XX
OS  Synecchocystis sp.
XX
PN  WO9904622-A1.
XX
PD  04-FEB-1999.
XX
PF  22-JUL-1998; 98WO-US15137.
XX
PR  17-JUL-1998; 98US-0053819.
PR  25-JUL-1997; 97US-0053819.
PR  26-JAN-1998; 98US-0072497.
XX
PA  (UYNE-) UNIV NEVADA.
XX
PI  Dellapenna D, Shintani DK;
XX
DR  WPI: 1999-142458/12.
DR  N-PSDB: AAX17788.
XX
XX  Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
PT  methyltransferase coding sequence - useful for producing
PT  ^a-tocopherol, and transgenic plants, seeds and oils with an altered
PT  tocopherol profile
XX
XX  Example 2: Page 33-35; 46pp; English.
XX
XX  The invention provides DNA sequences encoding gamma-tocopherol
CC  methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089
CC  and 165H5T7 are isolated from Synecchocystis and Arabidopsis species
CC  respectively. The DNA fragments are useful for producing transgenic
CC  plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
CC  are also useful for producing alpha-tocopherol and plants with increased
CC  gamma-tocopherol, which may be useful in certain industries such as the
CC  meat industry e.g. for developing forage plants to feed animals. The
CC  production of transgenic plants (and seeds) with a higher alpha-
CC  tocopherol:gamma-tocopherol ratio, using the gamma TMTs, will increase
CC  the level of alpha-tocopherol in the human diet, in addition to enhancing
CC  the stability and shelf life of plants and plant products. Increased
CC  levels of alpha-tocopherol will also increase meat quality and extend
CC  shelf life of post-processed meat products. Plants with a higher alpha-
CC  tocopherol:gamma-tocopherol ratio may also produce advantageous
CC  phenotypes. The present sequence represents a Synecchocystis gamma-TMT.
XX
XX
SQ  Sequence 317 AA:

Query Match          7.9%; Score 204.5; DB 20; Length 317;
Best Local Similarity 26.3%; Pred. No. 8.2e-11;
Matches 61; Conservative 48; Mismatches 90; Indels 33; Gaps 8;

OY  215 YVSKKKNQNIQSWLQKQVDEDDKGFQRFLLDSSQYKENSILKRVFPGYVSTGLET 274
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  35 yeklnfyddssglwedwge-----hmhgyyphngtyidr-----rqqlidl 80
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  275 KEF-----VSKLDKRGOKVLDVGGIGGDFYMAENVDEVWIDLSNMISFALE--RS 328
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  81 kellawapqnsakp-rxildgicgissjlylaqghaewngaslsipqveragera 139
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  329 IGLKCAVEFVADCTKKQYVENSFDVIVSRDILIHIOCKPALFRFHHMLKRGKGVLLSD 388
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  140 Iglstcgyvnaalldlpfasdsfivwvslsgehmpnkagllgawvllkpgglllat 199
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  369 YC-----KSAGTPEAEFAIYTRQGYDLH-----DVKAYGMLMDAGFEVY 429
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  200 wchpripdgngpltdaterrhl-gaiydyvclyplvvsldyealarecgfge 250
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT  4
AAV39308

```

```

ID  AAV39308 standard; Protein; 283 AA.
XX
AC  AAV39308;
XX
DT  01-DEC-1999 (first entry)
XX
DE  SpnL protein involved in spinosyn biosynthesis.
XX
KW  Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
KW  microtides; arachnid; nematode; insect; polyketide; polyketide synthase;
KW  PKS; extender module; initiator module; acyl transferase domain; A1;
KW  acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
KW  dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
KW  insecticide; C-methylation.
XX
OS  Saccharopolyspora spinosa.
XX
PN  WO9946387-A1.
XX
PD  16-SEP-1999.
XX
PF  16-FEB-1999; 99WO-US03212.
XX
PR  09-MAR-1998; 98US-0036987.
XX
PA  (DOMC ) DOW AGROSCIENCES LLC.
XX
PI  Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
PI  Treadway PJ, Turner JR, Waldron C;
XX
DR  WPI: 1999-551414/46.
DR  N-PSDB: AAZ21501.
XX
XX  New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
PT  for production of insecticidal spinosyn compounds
XX
XX  Claim 1; Page 156-157; 190pp; English.
XX
XX  This is the amino acid sequence of the product of the spnL gene. The
CC  protein is involved in spinosyn biosynthesis. The spnL gene is one of 23
CC  genes and open reading frames contained in an 80kb DNA sequence
CC  AA021501. Spinosyns are insecticidal microtides which are useful for the
CC  control of arachnids, nematodes and insects. Biosynthesis of spinosyns
CC  occurs via stepwise condensation and modification of carboxylic acid
CC  precursors generating a linear polyketide which is modified further. The
CC  DNA sequence contains a central region of approximately 55kb which has
CC  homology to the DNA encoding the polyketide synthases (PKS) of known
CC  macrolide producers. The spinosyn PKS DNA region consists of 5 ORFs with
CC  stop codons at the end of acyl carrier protein (ACP) domains. Together
CC  the PKS polypeptides (AAV39297-Y39301), form a complex consisting of an
CC  initiator module, spnA, and several extender modules spnB-spnE. The
CC  products of the genes present in the region upstream of the PKS genes
CC  have been assigned names spnF-spnS AAV39302-Y39315 and are responsible
CC  for different modifications in spinosyn biosynthesis. There are also two
CC  ORFs ORF15 and ORF16 present immediately upstream of spnS, producing
CC  polypeptides AAV39316-Y39317, and two ORFs ORF1 and ORF2 present
CC  downstream of the PKS region producing polypeptides AAV39318-Y39319. It
CC  is suggested that SpnL is involved in C-methylation during spinosyn
CC  biosynthesis. The genes are useful to improve yields of spinosyns and
CC  for creating new spinosyns e.g. by mutagenesis, or interruption of steps
CC  in spinosyn biosynthesis. The modified spinosyns may be a new insect
CC  control agent or serve as substrates for further chemical modification
CC  and the creation of new semi-synthetic spinosyns. The genes are also
CC  useful to isolate similar sequences from S. spinosa or other species by
XX  hybridization.
XX
SQ  Sequence 283 AA:

Query Match          7.6%; Score 198; DB 20; Length 283;
Best Local Similarity 25.7%; Pred. No. 2.8e-10;
Matches 53; Conservative 44; Mismatches 97; Indels 12; Gaps 5;

```

[illegible][illegible]

CC spanning 55kb of DNA. The invention includes an expression cassette
 CC comprising a mitomycin biosynthetic gene operably linked to a promoter,
 CC and host cells transformed with the cassette. The nucleotide, and protein
 CC sequences and the transformed host cells of the invention result in
 CC antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and
 CC antibiologic activities. The nucleotide sequences are used to elucidate the
 CC molecular basis for the biosynthesis of the mitosane ring system, as well
 CC as to engineer the biosynthesis of novel natural products, e.g.
 CC antibiotics, anti-inflammatory agents, anti-cancer agents,
 CC immune-enhancers, immunosuppressants, agents to treat asthma, chronic
 CC obstructive pulmonary disease as well as other disease involving
 CC respiratory inflammation, or cholesterol-lowering agents or as crop
 CC protection agents (e.g. fungicides or insecticides) as well as
 CC biopolymers, e.g., in packaging or biomedical applications, or to engineer
 CC PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and
 CC AAC32485-B32542 represent mitomycin biosynthetic gene cluster DNA
 CC sequences and encoded proteins. Sequences AAC55812-C55814,
 CC AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the
 CC cloning of the mitomycin biosynthetic genes.

XX Sequence 275 AA;

Query Match 7.3%; Score 189.5; DB 21; Length 275;
 Best Local Similarity 27.5%; Pred. No. 1.7e-09;
 Matches 61; Conservative 38; Mismatches 102; Indels 21; Gaps 7;

QY 263 PGYSTGLETKEFVSLDKPGOKVLDVGGCGDFYMAENVDEWGDINNMS 322

DB 41 pttveeaadrltdysetrlrpagervldvsgngkatrlraargraty--slnpqy 98

QY 323 FALERSTLAK--CAVEFEVADCTKKDYPENSFDVYSRDTILHIOKPALEFRSFHKWLK 379

DB 99 vglstqlaekgdeetetrldmalpfpdgsfdacysaleschaleradvfletarlr 158

QY 360 PEGKVLIDY-----CKSAGTPSAEFAAYIRORYDLHDVAKYGMKLDAG-----FVEV 429

DB 159 pgyrtvtdlrlrpldsarltvtdandrfqgspvl--treayedcmrsvglevefidi 217

QY 430 IAEKRTDOFIOVLOKELDALBEKEDFDIDFSEEDYNIDVGD 471

DB 218 gdevrps--yeavaakmrardelsglmd--eafhmrvdg 254

RESULT 7

AAB32507
 ID AAB32507 standard; Protein: 283 AA.

AC AAB32507;

DT 19-JAN-2001 (first entry)

DE S. lavendulae Mit M encoded protein sequence.

XX Mitomycin; biosynthesis; mitosane ring system; antibiologic; anti-cancer;

KW anti-inflammatory; immune-enhancer; immunosuppressant; asthma;

KW chronic obstructive pulmonary disease; respiratory inflammation;

KW fungicide; pesticide.

OS Streptomyces lavendulae.

XX WO200053737-A2.

XX 14-SEP-2000.

XX 10-MAR-2000; 2000WO-US06394.

XX 12-MAR-1999; 99US-0266965.

XX (MINU) UNIV MINNESOTA.

XX (SHER/) SHERMAN D H.

XX (MAOY/) MAO Y.

XX (VARO/) VAROGLU M.

PA (HEMU/) HE M.
 PA (SHEL/) SHELDON P C.
 XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;
 PI WPI: 2000-601980/57.
 DR N-PSDB: AAC55802.

PT Novel nucleic acid molecule comprising mitomycin biosynthetic gene
 PT cluster useful for cloning mitomycin biosynthetic genes for elucidating
 PT the molecular basis of mitosane ring system biosynthesis

PS Disclosure: Page 353; 399pp: English.

CC This invention relates to isolated and purified nucleic acid molecules
 CC from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
 CC natural products that contain a variety of functional groups, including
 CC amino benzoxazinone and axiridine ring systems. The S. lavendulae
 CC mitomycin biosynthetic gene cluster comprises 47 mitomycin genes
 CC spanning 55kb of DNA. The invention includes an expression cassette
 CC comprising a mitomycin biosynthetic gene operably linked to a promoter,
 CC and host cells transformed with the cassette. The nucleotide, and protein
 CC sequences and the transformed host cells of the invention result in
 CC antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and
 CC antibiologic activities. The nucleotide sequences are used to elucidate the
 CC molecular basis for the biosynthesis of the mitosane ring system, as well
 CC as to engineer the biosynthesis of novel natural products, e.g.
 CC antibiotics, anti-inflammatory agents, anti-cancer agents,
 CC immune-enhancers, immunosuppressants, agents to treat asthma, chronic
 CC obstructive pulmonary disease as well as other disease involving
 CC respiratory inflammation, or cholesterol-lowering agents or as crop
 CC protection agents (e.g. fungicides or insecticides) as well as
 CC biopolymers, e.g., in packaging or biomedical applications, or to engineer
 CC PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and
 CC AAC32485-B32542 represent mitomycin biosynthetic gene cluster DNA
 CC sequences and encoded proteins. Sequences AAC55812-C55814,
 CC AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the
 CC cloning of the mitomycin biosynthetic genes.

XX Sequence 283 AA;

Query Match 7.3%; Score 189.5; DB 21; Length 283;
 Best Local Similarity 25.4%; Pred. No. 1.8e-09;
 Matches 60; Conservative 52; Mismatches 93; Indels 31; Gaps 9;

QY 274 TKEFVSKLDIKRGOKVLDVGGCGDFYMAENVDEWGDLSINMT--SFALERSTGL 331

DB 60 tdmmaeerlrigasrvldlqgcvgfppvyrlasgahvtglsvshegyvranaaleea 119

QY 332 KCAVEFEVADCTKKDYPENSFDVYSRDTILHIOKPALEFRSFHKWLKPGKVLISDYCK 391

DB 120 adrarfrgradndlpfedcsdvalalesilmprragvlaqgvrlppgglvltldffe 179

QY 392 SAGTPSAEFAAYIRORYDLHDVAKYGMKLDAG-----FVEVIAENRTDQFI 439

DB 180 raplapegraaavgry---lhdftmtmwsaeayppllrigagjwleefld-lsdqtlekf 234

QY 440 QVLOKELDALBEKEDFDIDFSEEDYNIDVGMKAKIVRTTGEQOMG--LFIARK 493

DB 235 rliserlinskgfgle---tqfgeemvngfdpg---dlv---gvkefgylllvagr 280

RESULT 8

AAG21652
 ID AAG21652 standard; Protein: 293 AA.

AC AAG21652;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 24284.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
Arabidopsis thaliana.
EPI033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-0301439.
25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123380.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 05-MAY-1999; 990S-0132484.
PR 06-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135324.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.

PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 02-JUL-1999; 990S-0142154.
PR 06-JUL-1999; 990S-0142055.
PR 08-JUL-1999; 990S-0142390.
PR 09-JUL-1999; 990S-0142803.
PR 12-JUL-1999; 990S-0142920.
PR 13-JUL-1999; 990S-0142977.
PR 14-JUL-1999; 990S-0143542.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 26-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156536.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.28; Score 188; DB 21; Length 293;
Best Local Similarity 27.3%; Pred. No. 2.6e-09;

Matches 65; Conservative 36; Mismatches 93; Indels 44; Gaps 7;

OY 280 KLIDKPGOKVLDVGGGCGSGFYNAENVYVIGDLSINMISFA--LERSIGLKCAVEF 337
DB 46 q1g1qpgqkvdvqg1qgplre1arf1nsavtqlnneyq1ltgkelnrlagvdkcnf 105
OY 338 EVADCTKDYDENSFDVYSRDTILHIDQKPALEFRSPFKWLKPGCGKVLISDYCKS-AGTP 396
DB 106 vkadftmkmpfensdvdya1eatchapdaygcylryv1kpgqcfayekwcmtdafcp 165
OY 397 SAEFRAVYR---QKGYDLHDYKAYGK---MLKDAGFVYIAENR----- 434
DB 166 dnaehqk1gse1e1gdyg1pdir1ltck1eal1kgagf-ev1ewek1adkdsppwpy1pldkn 224
OY 435 -----TQGFIOV1QKELD1EQEKDPIIDPESSEDDYVDYDGMKAKL 476
DB 225 hfs1ss1fl1tavgrf1tknmvkl1--ey1rlapqsg1rvsn1l1eqaaeg1vdggr1rel 280

RESULT 9
AAG21526
ID AAG21526 standard; Protein: 296 AA.
XX
AC AAG21526;

XX XX
DT 17-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24113.
XX XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PE 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 17-JUN-1999; 99US-0139453.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.
XX PR 18-JUN-1999; 99US-0139463.

RESULT 10
AAG21651
ID AAG21651 standard; Protein; 336 AA.
XX
AC AAG21651;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24283.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123380.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

	PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
	PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
	PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
	PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
	PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
	PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
	PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
	PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
	PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
	PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
	PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
	PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
	PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151080.
	PR	23-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
	PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151303.
	PR	24-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
	PR	28-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
	PR	29-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
	PR	30-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
	PR	01-JUL-1999;	PR	15-SEP-1999;	99US-0154018.
	PR	01-JUL-1999;	PR	16-SEP-1999;	99US-0154039.
	PR	02-JUL-1999;	PR	20-SEP-1999;	99US-0154779.
	PR	06-JUL-1999;	PR	22-SEP-1999;	99US-0155139.
	PR	08-JUL-1999;	PR	23-SEP-1999;	99US-0155486.
	PR	09-JUL-1999;	PR	24-SEP-1999;	99US-0155659.
	PR	12-JUL-1999;	PR	28-SEP-1999;	99US-0156458.
	PR	13-JUL-1999;	PR	29-SEP-1999;	99US-0156596.
	PR	14-JUL-1999;	PR	04-OCT-1999;	99US-0157117.
	PR	15-JUL-1999;	PR	05-OCT-1999;	99US-0157753.
	PR	16-JUL-1999;	PR	06-OCT-1999;	99US-0157865.
	PR	16-JUL-1999;	PR	07-OCT-1999;	99US-0158029.
	PR	19-JUL-1999;	PR	08-OCT-1999;	99US-0158232.
	PR	19-JUL-1999;	PR	12-OCT-1999;	99US-0158369.
	PR	19-JUL-1999;	PR	13-OCT-1999;	99US-0159293.
	PR	19-JUL-1999;	PR	13-OCT-1999;	99US-0159284.
	PR	19-JUL-1999;	PR	13-OCT-1999;	99US-0159285.
	PR	19-JUL-1999;	PR	14-OCT-1999;	99US-0159295.
	PR	20-JUL-1999;	PR	14-OCT-1999;	99US-0159329.
	PR	20-JUL-1999;	PR	14-OCT-1999;	99US-0159330.
	PR	20-JUL-1999;	PR	14-OCT-1999;	99US-0159637.
	PR	21-JUL-1999;	PR	14-OCT-1999;	99US-0159637.
	PR	21-JUL-1999;	PR	14-OCT-1999;	99US-0159638.
	PR	21-JUL-1999;	PR	18-OCT-1999;	99US-0159584.
	PR	21-JUL-1999;	PR	21-OCT-1999;	99US-0160741.
	PR	22-JUL-1999;	PR	21-OCT-1999;	99US-0160767.
	PR	22-JUL-1999;	PR	21-OCT-1999;	99US-0160768.
	PR	22-JUL-1999;	PR	21-OCT-1999;	99US-0160770.
	PR	22-JUL-1999;	PR	21-OCT-1999;	99US-0160814.
	PR	22-JUL-1999;	PR	21-OCT-1999;	99US-0159637.
	PR	23-JUL-1999;	PR	21-OCT-1999;	99US-0160815.
	PR	23-JUL-1999;	PR	22-OCT-1999;	99US-0160980.
	PR	23-JUL-1999;	PR	22-OCT-1999;	99US-0160981.
	PR	26-JUL-1999;	PR	22-OCT-1999;	99US-0160989.
	PR	27-JUL-1999;	PR	25-OCT-1999;	99US-0161404.
	PR	27-JUL-1999;	PR	25-OCT-1999;	99US-0161405.
	PR	27-JUL-1999;	PR	25-OCT-1999;	99US-0161406.
	PR	27-JUL-1999;	PR	26-OCT-1999;	99US-0161359.
	PR	28-JUL-1999;	PR	26-OCT-1999;	99US-0161360.
	PR	02-AUG-1999;	PR	26-OCT-1999;	99US-0161361.
	PR	02-AUG-1999;	PR	28-OCT-1999;	99US-0161920.
	PR	03-AUG-1999;	PR	28-OCT-1999;	99US-0161920.
	PR	04-AUG-1999;	PR	28-OCT-1999;	99US-0161993.
	PR	04-AUG-1999;	PR	29-OCT-1999;	99US-0162142.
	PR	05-AUG-1999;	PR		
	PR	05-AUG-1			

PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148564.
PR	14-AUG-1999;	99US-0148684.
PR	15-AUG-1999;	99US-0149368.
PR	16-AUG-1999;	99US-0149175.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160989.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Dd		111 qtiqpgqkvldvgvsgigbpjrlrelarfensavtglmnegyitrgkclnrlagvdkcnf 170
Oy		338 EVADCTKKDYPENSEFDVIYSRDTIIHIDKPALFRSEFKMLKPGCKVLISDYCKS-AGTP 396 : : : : : : : : Dd 171 vkadfmknmpifpensfdavaylaatchapdaycykelrvtlpgqcfaayewcmtdafdp 230
Oy		397 SAEFAAYIR---QRRYVDLHDVKAQKC---MKKDGFVEVTIENR-----434 : : : : : Dd 231 dnaehqklkgelelygdgipdlriltckclealkgaqf-evlewelakdsppwylpdkn 289
Oy	435 -----	-TQOIFIVLOKELDALOEKDDFIDEFSEEDYNNDIVDGWRKL 476 - : : : : : : : : : Dd 290 hfslsfrltavgrrtlkmwkiil--eyirlapgagsqrsvnfleqaagvlvdggrrci 345
RESULT	13	
AACG21650		
ID	AACG21650 standard; Protein: 366 AA.	
XX	AACG21650;	
XX		
DT	17-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24282.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KM	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PE	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	
PR	23-APR-1999; 99US-0130510.	
PR	28-APR-1999; 99US-0130891.	
PR	30-APR-1999; 99US-0131449.	
PR	30-APR-1999; 99US-0132048.	
PR	30-APR-1999; 99US-0132407.	
PR	04-MAY-1999; 99US-0132484.	
PR	05-MAY-1999; 99US-0132485.	
PR	06-MAY-1999; 99US-0132486.	
PR	07-MAY-1999; 99US-0132487.	
PR	11-MAY-1999; 99US-0132863.	
PR	14-MAY-1999; 99US-0134256.	
PR	14-MAY-1999; 99US-0134218.	
PR	14-MAY-1999; 99US-0134219.	
PR	14-MAY-1999; 99US-0134221.	
PR	14-MAY-1999; 99US-0134370.	
PR	18-MAY-1999; 99US-0134376.	
PR	19-MAY-1999; 99US-0134941.	
PR	20-MAY-1999; 99US-0135124.	
PR	21-MAY-1999; 99US-0135353.	
PR	24-MAY-1999; 99US-0135629.	
PR	25-MAY-1999; 99US-0136021.	
PR	27-MAY-1999; 99US-0136392.	
PR	28-MAY-1999; 99US-0136782.	
PR	01-JUN-1999; 99US-0137222.	

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148665.
PR 16-AUG-1999; 99US-0148668.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0150885.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154409.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159320.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-01610814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

7.2%; Score 188; DB 21; Length 366;

Best Local similarity 27.3%, Pred. No. 3 6e-09; Matches 65; Conservative 36; Mismatches 93; Indels 44; Gaps 7;

OY	280	KLDIKPGGKVLIDVCGGTGGDPFYAENDEVVEYGIDLSINMISFA--LERSTGLCAVEF	337
Dd	119	: : : : : : : : : : : : : : :	178
OY	338	EVAACCTKKDYDENSHDVYSRDTILHIDDKALFEFSFKWKLPGGVKLVSDYCKS-AAGP	396
Dd	179	vkdafimkmpfensfidavayaieatchapdaygyeylrvklpgqcafaayewcmctardcp	238
OY	397	SAEFAAYR---QGCYDHDHKAVKC---MKKDGFVVILENR-----	434
Dd	239	dnaehgikgetelvgdlipdrilrttkclealkgaif-eiwekldiakspypwpylipdkn	297
OY	435	-----TDQTQVLOKEIDLAEQEKKDFIDFSEEDYNDVIDGWAKTL	476
Dd	298	nfsiasftltavgfrilknmwkll--eqlrltlaaggsgqvsnilegaanaeglvdggrtel	353

RESULT 14
AAG45952
ID AAG45952 standard; Protein; 293 AA.

AC AAG45952;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 57754.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter:::
KW termination sequence.

05 *Arabidopsis thaliana*.

PN EP1033405-A2

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR	2-FEB-1999	99US-0121825
PR	05-MAR-1999	99US-0123180
PR	09-MAR-1999	99US-0123548
PR	23-MAR-1999	99US-0125764
PR	25-MAR-1999	99US-0126586
PR	29-MAR-1999	99US-0126785
PR	01-APR-1999	99US-0127462
PR	06-APR-1999	99US-0128234
PR	08-APR-1999	99US-0128714
PR	16-APR-1999	99US-0129845
PR	19-APR-1999	99US-0130047
PR	21-APR-1999	99US-0130049
PR	23-APR-1999	99US-0130510
PR	23-APR-1999	99US-0130891
PR	28-APR-1999	99US-0131449
PR	30-APR-1999	99US-0132048
PR	30-APR-1999	99US-0132407
PR	04-MAY-1999	99US-0132485
PR	05-MAY-1999	99US-0132486
PR	06-MAY-1999	99US-0132486
PR	06-MAY-1999	99US-0132487
PR	07-MAY-1999	99US-0132863
PR	11-MAY-1999	99US-0134256
PR	14-MAY-1999	99US-0134219
PR	14-MAY-1999	99US-0134219
PR	14-MAY-1999	99US-0134221
PR	14-MAY-1999	99US-0134370
PR	18-MAY-1999	99US-0134768
PR	19-MAY-1999	99US-0134941
PR	20-MAY-1999	99US-0135124
PR	21-MAY-1999	99US-0135523

PR	24-MAY-1999	9905-01356529
PR	25-MAY-1999	9905-01360621
PR	27-MAY-1999	9905-01363692
PR	28-MAY-1999	9905-01367822
PR	01-JUN-1999	9905-01372722
PR	03-JUN-1999	9905-01375728
PR	04-JUN-1999	9905-01375702
PR	07-JUN-1999	9905-01377274
PR	08-JUN-1999	9905-01380094
PR	10-JUN-1999	9905-01385400
PR	10-JUN-1999	9905-01388147
PR	14-JUN-1999	9905-01391119
PR	16-JUN-1999	9905-01394552
PR	16-JUN-1999	9905-01394552
PR	16-JUN-1999	9905-01394552
PR	17-JUN-1999	9905-01394932
PR	18-JUN-1999	9905-01394954
PR	18-JUN-1999	9905-01394954
PR	18-JUN-1999	9905-01394962
PR	18-JUN-1999	9905-01394663
PR	18-JUN-1999	9905-01394557
PR	18-JUN-1999	9905-01394559
PR	18-JUN-1999	9905-01394660
PR	18-JUN-1999	9905-01394661
PR	18-JUN-1999	9905-01394662
PR	18-JUN-1999	9905-01394662
PR	18-JUN-1999	9905-01394663
PR	18-JUN-1999	9905-01396117
PR	21-JUN-1999	9905-01396199
PR	22-JUN-1999	9905-01403553
PR	22-JUN-1999	9905-01403553
PR	23-JUN-1999	9905-01403554
PR	24-JUN-1999	9905-01405695
PR	24-JUN-1999	9905-01408923
PR	28-JUN-1999	9905-01409821
PR	30-JUN-1999	9905-01412877
PR	01-JUL-1999	9905-01418342
PR	01-JUL-1999	9905-01421554
PR	02-JUL-1999	9905-01420055
PR	06-JUL-1999	9905-01423900
PR	08-JUL-1999	9905-01428203
PR	09-JUL-1999	9905-01442820
PR	12-JUL-1999	9905-01428270
PR	13-JUL-1999	9905-01435324
PR	14-JUL-1999	9905-01435324
PR	15-JUL-1999	9905-01440085
PR	16-JUL-1999	9905-01440085
PR	16-JUL-1999	9905-01440866
PR	19-JUL-1999	9905-01443325
PR	19-JUL-1999	9905-01443325
PR	19-JUL-1999	9905-01443324
PR	19-JUL-1999	9905-01443332
PR	19-JUL-1999	9905-01443332
PR	21-JUL-1999	9905-01448114
PR	21-JUL-1999	9905-01450866
PR	21-JUL-1999	9905-01450888
PR	22-JUL-1999	9905-01450887
PR	22-JUL-1999	9905-01450889
PR	22-JUL-1999	9905-01450899
PR	22-JUL-1999	9905-01451545
PR	23-JUL-1999	9905-01451518
PR	23-JUL-1999	9905-01455224
PR	26-JUL-1999	9905-01455776
PR	27-JUL-1999	9905-01455713
PR	27-JUL-1999	9905-01455919
PR	28-JUL-1999	9905-01459511
PR	02-AUG-1999	9905-01463686
PR	02-AUG-1999	9905-01465888

PR	02-AUG-1999;	9905-0146388;
PR	03-AUG-1999;	9905-0147038
PR	04-AUG-1999;	9905-0147204
PR	04-AUG-1999;	9905-0147302
PR	05-AUG-1999;	9905-0147392
PR	05-AUG-1999;	9905-0147260
PR	06-AUG-1999;	9905-0147303
PR	06-AUG-1999;	9905-0147416
PR	09-AUG-1999;	9905-0147493
PR	09-AUG-1999;	9905-0147935
PR	10-AUG-1999;	9905-0148171
PR	11-AUG-1999;	9905-0148319
PR	12-AUG-1999;	9905-0148341
PR	13-AUG-1999;	9905-0148565
PR	13-AUG-1999;	9905-0148684
PR	16-AUG-1999;	9905-0149368
PR	17-AUG-1999;	9905-0149175
PR	18-AUG-1999;	9905-0149426
PR	20-AUG-1999;	9905-0149722
PR	20-AUG-1999;	9905-0149723
PR	23-AUG-1999;	9905-0149920
PR	23-AUG-1999;	9905-0149930
PR	25-AUG-1999;	9905-0150566
PR	26-AUG-1999;	9905-0150884
PR	27-AUG-1999;	9905-0151065
PR	27-AUG-1999;	9905-0151066
PR	27-AUG-1999;	9905-0151080
PR	30-AUG-1999;	9905-0151303
PR	31-AUG-1999;	9905-0151348
PR	01-SEP-1999;	9905-0151930
PR	07-SEP-1999;	9905-0152363
PR	10-SEP-1999;	9905-0153070
PR	13-SEP-1999;	9905-0153758
PR	15-SEP-1999;	9905-0154018
PR	16-SEP-1999;	9905-0154039
PR	20-SEP-1999;	9905-0154779
PR	22-SEP-1999;	9905-0155139
PR	23-SEP-1999;	9905-0155486
PR	24-SEP-1999;	9905-0155569
PR	28-SEP-1999;	9905-0156458
PR	29-SEP-1999;	9905-0156566
PR	04-OCT-1999;	9905-0157117
PR	05-OCT-1999;	9905-0157753
PR	06-OCT-1999;	9905-0157865
PR	07-OCT-1999;	9905-0158029
PR	08-OCT-1999;	9905-0158232
PR	12-OCT-1999;	9905-0158369
PR	13-OCT-1999;	9905-0158529
PR	13-OCT-1999;	9905-0158294
PR	13-OCT-1999;	9905-0158525
PR	14-OCT-1999;	9905-0158329
PR	14-OCT-1999;	9905-0155330
PR	14-OCT-1999;	9905-0155331
PR	14-OCT-1999;	9905-0159637
PR	14-OCT-1999;	9905-0159638
PR	21-OCT-1999;	9905-0160814
PR	21-OCT-1999;	9905-0160845
PR	22-OCT-1999;	9905-0160980
PR	22-OCT-1999;	9905-0160981
PR	22-OCT-1999;	9905-0160989
PR	25-OCT-1999;	9905-0161404
PR	25-OCT-1999;	9905-0161405
PR	25-OCT-1999;	9905-0161406
PR	26-OCT-1999;	9905-0161359
PR	26-OCT-1999;	9905-0161360
PR	28-OCT-1999;	9905-0161361
PR	28-OCT-1999;	9905-0161920

[illegible]

RESULT	15
AA645955	
ID	AA645955 standard; Protein; 293 AA.
XX	
AC	AA645955;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 57758.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PT	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	15-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136332.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138004.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:16:42 : Search time 13.12 Seconds
(Without alignments)
775.277 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598
Sequence: 1 MAASAMGVLOEREVFKKKYMI.....KIVRTTEGQGMFLIAKKM 494

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	207	8.0	363	4 US-09-041-718-5	Sequence 5, App1
2	198	7.6	283	4 US-09-036-987A-13	Sequence 13, App1
3	179	6.9	280	4 US-09-029-603-2	Sequence 2, App1
4	172.5	6.6	282	1 US-08-457-245-9	Sequence 9, App1
5	170	6.5	383	4 US-09-041-718-3	Sequence 3, App1
6	168	6.5	275	4 US-09-036-987A-7	Sequence 7, App1
7	160	6.2	361	4 US-09-041-718-4	Sequence 4, App1
8	148.5	5.7	376	4 US-09-041-718-2	Sequence 2, App1
9	128.5	4.9	276	3 US-08-935-263-16	Sequence 16, App1
10	128.5	4.9	308	1 US-08-457-245-8	Sequence 8, App1
11	122.5	4.7	1086	6 5386025-8	Patent No. 5386025
12	120	4.6	1084	1 US-08-455-543A-56	Sequence 56, App1
13	120	4.6	1084	2 US-08-223-305C-56	Sequence 56, App1
14	120	4.6	1103	1 US-08-455-543A-53	Sequence 53, App1
15	120	4.6	1103	2 US-08-223-305C-53	Sequence 25, App1
16	115	4.4	287	1 US-08-457-245-3	Sequence 3, App1
17	114.5	4.4	1091	1 US-07-745-206A-25	Sequence 25, App1
18	114.5	4.4	1091	1 US-08-455-543A-52	Sequence 52, App1
19	114.5	4.4	1091	2 US-08-223-305C-52	Sequence 52, App1
20	114.5	4.4	1091	2 US-08-311-363-25	Sequence 25, App1
21	114.5	4.4	1091	3 US-08-713-118-4	Sequence 4, App1
22	114.5	4.4	1091	4 US-09-452-007-4	Sequence 4, App1
23	114	4.4	1079	1 US-08-455-543A-55	Sequence 55, App1
24	114	4.4	1079	2 US-08-223-305C-55	Sequence 55, App1
25	113.5	4.4	1106	1 US-08-435-675B-5	Sequence 5, App1
26	110.5	4.3	1106	1 US-08-336-257A-8	Sequence 8, App1
27	108.5	4.2	1086	1 US-08-455-543A-54	Sequence 54, App1

28	108.5	4.2	1086	2 US-08-223-305C-54	Sequence 54, App1
29	105	4.0	715	2 US-08-484-993B-10	Sequence 10, App1
30	105	4.0	715	2 US-08-484-158B-10	Sequence 10, App1
31	105	4.0	715	2 US-08-480-596A-10	Sequence 10, App1
32	105	4.0	715	2 US-08-480-150A-10	Sequence 10, App1
33	105	4.0	715	3 US-08-458-731-10	Sequence 10, App1
34	105	4.0	715	3 US-08-149-223A-10	Sequence 10, App1
35	104.5	4.0	715	3 US-08-361-083-64	Sequence 64, App1
36	102	3.9	773	2 US-08-966-389-4	Sequence 4, App1
37	102	3.9	773	2 US-09-103-509-4	Sequence 4, App1
38	102	3.9	773	2 US-09-102-644-4	Sequence 4, App1
39	102	3.9	773	2 US-09-218-032-4	Sequence 4, App1
40	100	3.8	460	3 US-09-147-522-9	Sequence 9, App1
41	99.5	3.8	600	1 US-08-374-155A-10	Sequence 10, App1
42	99.5	3.8	600	2 US-08-785-396-10	Sequence 10, App1
43	99	3.8	629	1 US-08-374-155A-4	Sequence 4, App1
44	99	3.8	629	2 US-08-785-396-4	Sequence 4, App1
45	99	3.8	974	3 US-08-938-291A-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-09-041-718-5
; Sequence 5, Application US/09041718A
; Patent No. 6225075
; GENERAL INFORMATION:
; APPLICANT: Bard, Martin
; TITLE OF INVENTION: DNA encoding sterol methyltransferase
; FILE REFERENCE: 740.003051
; CURRENT APPLICATION NUMBER: US/09/041,718A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Triticum atrivum
US-09-041-718-5

Query Match 8.0%; Score 207; DB 4; Length 363;
Best Local Similarity 28.6%; Pred. No. 3.9e-12;
Matches 68; Conservative 38; Mismatches 88; Indels 44; Gaps 8;

QY 280 KLDLPGKQVLDVGGCGIGGDFVMAENVYDEYVIGDLSIMWIS--FALERSIGKCAVEF 337
DB 114 QLELPGKQVLDVGGCGIGGDFVMAENVYDEYVIGDLSIMWIS--FALERSIGKCAVEF 173
QY 338 EVADCTKPDYNSFDVYISRDITLHIDQRPALFRSFHMLKPGKVLISDYC-KSAGTP 396
DB 174 VKADFMKMPSDNFDVAVYIEATCHAPDVGCGKELRYRLKPGQCFVAYEMCTIDHYP 233
QY 397 SAEPAAVIR---ORGYLDHVKAYGKML---KDAEVEVLT-----AENR----- 434
DB 234 NNATHKRIKDEIEIGNGIPRISTRCIOAVKDAE--EVIWDKDLADSDPLPWTLPDPS 292
QY 435 -----TDQFLVQKELDLAQKDPIDFSDREYNDIVDGKRAKL 476
DB 293 RFSLSRFLTTGVRITIRNMVKVL--EYVGAPRGSQSVSSFLKAAEGALVEGGKKEI 348
RESULT 2
US-09-036-987A-13
; Sequence 13, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madhuri, Krishnamurthy
; APPLICANT: Metto, Donald J.

```

; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes for Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-036-987A-13

```

Query Match 7.6%; Score 198; DB 4; Length 283;

Best Local Similarity 25.7%; Pred. No.2e-11;

Matches 53; Conservative 44; Mismatches 97; Indels 12; Gaps 5;

```

QY 272 ETTKEVSKLDKPGOKVLDVGGGIGGDFYMAENDVEVVGIDLSINMISFA--LERSI 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 38 QLDLFDIKALRRGALFLDLCGCGNGPVRACASGVRTGIVNQHIAATRLANET 97
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 GLCAVFEFVADCTKKDYSPNSFDVYSRDTILHIOCKPALFRSFHKWLKPGKVLISDY 389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 98 GLAGSLEFDLVDAQOLPYRPGFOAMAMOSVVOIVDOAAIREVHRILFPGGRFVIGDI 157
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 390 CKSAGTPSAEFAAYIRQGYDLHVKAYGKMLDAGF---VEVIAENR--TDQFIQVL 442
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 ITVRRLPE-EYAIV--WTGTATHTLNSFTALVSEAGFELLEVDLTNQTCMWSWYDEL 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 443 QKELDALEQKDFIDDESEEDYNDI 468
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 215 LRLKLDLAGVEPAVGYQORLYGDI 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 3

US-09-029-603-2

; Sequence 2, Application US/09029603

; Patent No. 6210935

; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas

; APPLICANT: Engel, Natalie

; APPLICANT: Bietenhader, Jurg

; APPLICANT: Toupet, Christine

; APPLICANT: Pospiech, Andreas

; TITLE OF INVENTION: staurosporin Biosynthesis Gene Clusters

; FILE REFERENCE: 4-20555/A/PCT

; CURRENT APPLICATION NUMBER: US/09/029,603

; CURRENT FILING DATE: 1998-03-20

```

; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; OTHER INFORMATION: methyl transferase-like protein
; US-09-029-603-2

```

Query Match 6.9%; Score 179; DB 4; Length 280;

Best Local Similarity 24.3%; Pred. No.1.5e-09;

Matches 49; Conservative 37; Mismatches 72; Indels 44; Gaps 6;

```

QY 257 YERV-----FGP---GYVSTGGLETTYKE-----FVSKLDKPGOKVLDVGC 294
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 18 YDRLTISAMNDGSFNPVNHIGYWDTPGSEATIEBAMDRLTDVTERLNAYATSHVLDLC 77
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 295 GIGGDFYMAENDVEVVGIDLSINMISFA--LERSIGLCAVEFEVADCTKKDYSPNSF 352
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 78 GVGGPGLRVARTGARVTGISISEQIRTNRLAAGVADRAVFGHGMKLPFADASP 137
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 DVYISRDTILHIOCKPALFRSFHKWLKPGKVLISD-----YCKSAGT 395
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 138 DAYMALESICHMPDROOVTEFVRCVRLRPGGRIVLTDLFFENPKAVRHRPIDKFCDLMS 197
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 PSA--EFAAYIRQGYDLHDV 414
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 198 TTDIDIDYVALLHRSGRLREI 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 4

US-08-457-245-9

; Sequence 9, Application US/08457245

; Patent No. 5573915

; GENERAL INFORMATION:

; APPLICANT: BARRY III, Clifton E.

; TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED

; TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Kourile and Crew

; STREET: Stewart Street Tower, One Market Plaza

; CITY: San Francisco

; STATE: California

; COUNTRY: US

; ZIP: 94105-1493

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/457,245

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Chambers, Guy W.

; REGISTRATION NUMBER: 30,617

; REFERENCE/DOCKET NUMBER: 15280-216000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 282 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

RESULT 13

```

0Y      77 KASQVATLDFIESYK-----KNESINCHYKNKVMC-----ADVTSPS 113
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```



```

Db 533 KSQEPVTLDFDAELENDIKYKIRNKMTDGSGETFRTLVKSODERYIDKGNRTYTWP 592
QY 116 LNISFNSVDIIFSWMILMYLSDSEVERLVE-RMLKMLKP-----GGYIFF--RESC----- 163
Db 593 VNGIDYSLALVLPYSPFYIKAKLEETITQARYSETLKPDPNFEESGYFFIAPRODCNDLK 652
QY 164 -----FHOSGDHKRSNPTHYREPPRYKIFKECHMODDSGNSY---ELSLIGC 209
Db 653 ISDNTEFLNFMNEFIDKRTNPNPSCNAD--LINRVLLDAGFTNELVONWYSKOKNIKGV 710
QY 210 KC-----IGAVYKSKKNOISWLMOK--VDSDDKGFQFRLDSSQYKFNSTIARYERV 260
Db 711 KARFVYTDGGITRYYPKAGEN-----WQENPEYEDSFYKRSLDNDNYVFTA--PYFNK 763
QY 261 FGPQVSTGCL--ETTKEFVSKLDLKPQOKVLDVCGGIGGDFYMAENVYDVEVVIDLSI 318
Db 764 SGPGAYESGIMVSKAVEIYIOGKLKPA-----VVGIRKIDV 799
QY 319 N-MISFALERSIGLKCAVEFEVADCTKDYSPNSFDVITYSDTILIHIDKPLFRSPFKW 377
Db 800 NSWIENFTKTSIRBPCA--GPVCDCKR-----NS-DVM--DCVI----- 833
QY 378 LKPGKVLIS--DYCKSAG-----TPS-----AEFAAYIRORGYDLHDYKAYGKMLKD 423
Db 834 LDDGFFILMANHDDYTNQIGRFEGEIDPSLMRHLVNISYVAFNKSYDQSCVCEPGAPKQ 893
QY 424 -----AGFEVIAE-----NRTDOFIQVL--QKELDALQEKDQDFIDFSEE 463
Db 894 GAGHRSAVPSVADILQIGMWATAAAMSILQOFLSLTFPRLLEAVEMEDDFTASLSKQ 953

```

```

APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1103 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-53

```

```

Query Match 4.6%; Score 120; DB 1; Length 1103;
Best Local Similarity 19.6%; Pred. No. 0.0077;
Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

```

```

RESULT 14
US-08-455-543A-53
Sequence 53, Application US/08455543A
Patent No. 5792846
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1103 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-53

Query Match 4.6%; Score 120; DB 1; Length 1103;
Best Local Similarity 19.6%; Pred. No. 0.0077;
Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

QY 77 KASQVIALDFIESVYK-----KNESINGHYKNKKEWC-----ADVSPS 115
Db 552 KSQEPVTLDFDAELENDIKYKIRNKMTDGSGETFRTLVKSODERYIDKGNRTYTWP 611
QY 116 LNISFNSVDIIFSWMILMYLSDSEVERLVE-RMLKMLKP-----GGYIFF--RESC----- 163
Db 612 VNGIDYSLALVLPYSPFYIKAKLEETITQARYSETLKPDPNFEESGYFFIAPRODCNDLK 671
QY 164 -----FHOSGDHKRSNPTHYREPPRYKIFKECHMODDSGNSY---ELSLIGC 209
Db 672 ISDNTEFLNFMNEFIDKRTNPNPSCNAD--LINRVLLDAGFTNELVONWYSKOKNIKGV 729
QY 210 KC-----IGAVYKSKKNOISWLMOK--VDSDDKGFQFRLDSSQYKFNSTIARYERV 260
Db 730 KARFVYTDGGITRYYPKAGEN-----WQENPEYEDSFYKRSLDNDNYVFTA--PYFNK 782
QY 261 FGPQVSTGCL--ETTKEFVSKLDLKPQOKVLDVCGGIGGDFYMAENVYDVEVVIDLSI 318
Db 783 SGPGAYESGIMVSKAVEIYIOGKLKPA-----VVGIRKIDV 818
QY 319 N-MISFALERSIGLKCAVEFEVADCTKDYSPNSFDVITYSDTILIHIDKPLFRSPFKW 377
Db 819 NSWIENFTKTSIRBPCA--GPVCDCKR-----NS-DVM--DCVI----- 852
QY 378 LKPGKVLIS--DYCKSAG-----TPS-----AEFAAYIRORGYDLHDYKAYGKMLKD 423
Db 853 LDDGFFILMANHDDYTNQIGRFEGEIDPSLMRHLVNISYVAFNKSYDQSCVCEPGAPKQ 912
QY 424 -----AGFEVIAE-----NRTDOFIQVL--QKELDALQEKDQDFIDFSEE 463
Db 913 GAGHRSAVPSVADILQIGMWATAAAMSILQOFLSLTFPRLLEAVEMEDDFTASLSKQ 972

RESULT 15
US-08-223-305C-53
Sequence 53, Application US/08223305C
Patent No. 5851824
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven

```

```

APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/866,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1103 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-223-305C-53

```

Query Match 4.6%; Score 120; DB 2; Length 1103;
 Best Local Similarity 19.6%; Pred. No. 0.0077;
 Matches 94; Conservative 76; Mismatches 158; Indels 152; Gaps 26;

```

QY 77 KASQVIALDFESYIK-----KNESINGHYKKNVKEMC-----ADVTSPS 115
DB 552 KSQEPVLDLDALELDKIKRNMKIDSGEKTFTLVKSODERYIDKGNFTYTWP 611
QY 116 LNI SPNSVDIIFSNWLMYLSDEEVERLVE-RMLKWLKP-----GGYIEF--RESC---- 163
DB 612 VNGIDYSLAVLVPYTSYYIAKLEBITTOARSETILKPDNFESGTYTIAFRDYCNDLK 671

```

```

QY 164 -----FHSGDHRKRSNPTHYRPRFYTKIFRECHMODDSGNSY---ELSLIGC 209
DB 672 ISDNNTFEFLNPNFEIDKKTNPNSCND--LIRVLLDAGFTNELVONYSKQKNIKGV 729
QY 210 KC-----IGATYKSKKNOISWLMQK-VSEEDKGFORFLDSSQYKNSILRTERV 260
DB 730 KAPVVTIDGITYRYPKEAGEN-----WQENPEYSEFSYRSISLDNDYVETA--PYFNK 782
QY 261 FGPVVSNGGL--ETTKFEVSKLDLKPQGVLDVCGGIGGSDFYMAENYDVEVYIDLSI 318
DB 783 SGPGAYESGIMYSKAVEIYIGKILKPA-----VGIKIDV 818
QY 319 N-MISFALERSIGLKCAVEFEVADCTKDPENSFDVLYSDTLIHLQDKPALERSFHKW 377
DB 819 NSWIEENFTKTSIRDPKA--GPVCDCKR-----NS-DVM--DCVI----- 852
QY 378 LKPGKVLIS--DYCKSAG-----TPS-----AEFAAYTRQGYDLHDVKAYGKMLKD 423
DB 853 LDDGGFLTMANHDDYTNQIGRPFGEIDPSLMRHLVNIISVAFNKSVDYQSVCEPGAAPKO 912
QY 424 -----AGEVEVIAE-----NRTDQFIQVL--QKELDALQEKDDFTIDFSEE 463
DB 913 GAGHSASVPSVADILQIGMMATAAAMSILQPFILSLFPRLLEAVENEDDFTASLSKQ 972

```

Search completed: September 20, 2001, 16:19:43
 Job time: 181 sec

```

Db      371  DVGIDLSVNMISFALBHAIGLKCSVEFEVADCTKEYPNDTFVYSRDTLHIOKPA 430
          :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY      370  LFRSFHMLKPGCKVLISDCKSAGTSPAFAAIRRGDLDHVKAYGKMLKAGFEV 429
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      431  LFRFXYMLKPGCKVLITDCCRSPKPSPDFAIYIKRGYDLHVOYGMRLPAGFEV 490
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      430  IAEKRTDQFQVLOKELDALEOEKDFIDFSEEDYNDIYDGWAKLVRTTEGQGLF 489
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      491  IAEKRTDQFQVLOKELDALEOEKDFIDFSEEDYNDIYDGWAKLVRTTEGQGLF 550
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      490  IAKK 493
          |||||
Db      551  IAKK 554

RESULT  2
F96525
protein T1N15.23 (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96525
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141, MGIID:21016719
A:Accession: F96525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <STO>
A:Cross-references: GB:AE005173; NID:g8778697; PIDN:AAF97905.1; GSPDB:GN00141
C:Genetics:
A:Gene: T1N15.23
A:Map position: 1

Query Match          49.6%; Score 1287.5; DB 2; Length 374;
Best Local Similarity 66.4%; Pred. No. 5.2e-83;
Matches 243; Conservative 36; Mismatches 28; Indels 59; Gaps 2;

OY      187  KIRKCHMDDSGNSYELSLIGCKICGAYKSKKNQNOISMOKYDSEDDKGFQRLDS 246
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      8  QVQECOTRNASGNSFELSMVGCKIGAYKKNQNOIMQKSVENDKDFQRLDN 67
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      247  SOKFNSILRYENFGYGTGGL-----ERTKE 276
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      68  VQKSSGILREYFGEYVSTGFGNSITLSSYGHITLYCLSYFMFLPSTETTK 127
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      277  FVSKDLKPGQKLVDCGCGIGGDFYMAENYDVEVGDLSINMISFALERSIGLKCAVE 336
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      128  FVAKMDLKPQKVLVDCGCGIGGDFYMAENYDVEVGDLSINMISFALERSIGLKCAVE 187
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      337  FEVADCTKRYPNSFDVYISRDTLHIOKPALEFRSFHMLKPGCKVLISDCKSAGT 396
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      188  FEVADCTKRYPNSFDVYISRDTLHIOKPALEFRSFHMLKPGCKVLITDCCRSAET 247
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      397  SAEFAAYIRRGYDLHVKAYGKMLKAGFEVYAEKRTDQFQVLOKELDALEOEKDF 456
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      248  SPFEAEIKRGYDLHVKAYGKMLKAGFEVYAEKRTDQFQVLOKELDALEOEKDF 307
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      457  IDDFSE-----EDYNDIYDGWAKLVRTTEGQGLF 487
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      308  IDDFSEYKFTLTNSFEQNTCLCSHIDKILLQEDYNDIYDGWAKLVRTTEGQGLF 367
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      488  LFTAK 493
          |||||

```

```

Db      368  LFTADK 373

RESULT  3
T27930
hypothetical protein F54D11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27930
R:Pauley, A.; Gattung, S.
Submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid F54D11.
A:Reference number: Z20606
A:Accession: T27930
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-437 <PAU>
A:Cross-references: EMBL:U64834; PIDN:AA04824.1; GSPDB:GN00023; CESP:F54D11.1
A:Experimental source: strain Bristol N2; clone F54D11
C:Genetics:
A:Gene: CESP:F54D11.1
A:Map position: 5
A:Introns: 36/2; 76/3; 108/1; 315/3; 398/3

Query Match          19.2%; Score 498.5; DB 2; Length 437;
Best Local Similarity 32.6%; Pred. No. 1.7e-27;
Matches 127; Conservative 70; Mismatches 156; Indels 37; Gaps 10;

OY      121  NSVDIIFSNMLWYLSDEVERLVERMLKMKP-----GYIFFRESCFHSGDHRK 173
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      64  NADVLIRNNALSOITINADL-----LTDFLKATNATATAGVITIED-LKDCSDKQV 116
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      174  SNPTHYREPRFYTKIFEKCHMODDSGNSYELSLIGCKCI--GAYVSKKNQNOISMW-- 229
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      117  ARLDY-----FVFRPTSDGNGNTGDIYVDQVHSNVE---QNFDFITVF 163
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      230  -QKV-----DSEDDKGFQRLDSQYKFNLSILRYERFPGYVSTGGLTTEFYSKL-DLK 284
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      164  RKRYFAPTDATTFRDLKQTYTNTGIDAYEMFVNFISPGYDENLKIKRFQDFK 223
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      285  PGQKVLVDCGCGIGGDFYMAENYDVEVGDLSINMISFALERSIGLK-CAVEVADCT 343
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      224  PGQTMLDIGVIGGAGARVADDFGVHGHGIDLSNMALALERLHEKDSRVKSTIDAL 283
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      344  KQDPENSPFVYISRDTLHIOKPALEFRSFHMLKPGCKVLISDCKSAGTSPAFAAY 403
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      284  VQFEDNSFYVFSRDTQIHPTTEKLSKIKYALKPGKVLITMYCKYGEQSDKRTY 343
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      404  IROGYDLHVKAYGKMLKAGFEVYAEKRTDQFQVLOKELDALEOEKDFIDFSEE 463
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      344  VQRAYFLKMLKEIADIANKTGFVNVQTEMTTPRFKILLEGHLEQNAERMSKRTQ 403
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY      464  DYNDIYDGWAKLVRTTEGQGLFIKK 493
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      404  ERDSLISGWTDKLYLEKDNHNNFFLAQ 433
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT  4
T27936
hypothetical protein ZK622.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T27936
R:Leinbach, D.
Submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid ZK622.
A:Reference number: Z20443
A:Accession: T27936
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-495 <LEI>
A:Cross-references: EMBL:U39998; PIDN:AA81102.1; CESP:ZK622.3

```



```

Db 155 VKADFMKMPEDNTFDVAVATEATCHAPDPVGCYKEIYRVLKPGCAFVEMCITDHYDP 214
      || : : || : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 397 SAEFAAIIIR---QNGYLDHVKAYGKML---KDAFVEV-----IAENR----- 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 NNATHKRIKDEIELGNGLPIRSTRQCLQAVKDAFGEVWMDKDLAEDSPLEWYLPDLPSR 274
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 435 -----TDQFIQVLQKELDALEQEKDFIDFSEEDYNDIYDGMKAKL 476
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 275 FSLSSFLTSVGRNITRTWKAL--EYVGLAPQGSERVSNTLEKRAEGIVEGKKEI 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

```

C96673
gamma-Glucosyltransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96673
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
  Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
  nansen, N.F.; Hughes, B.; Hultzer, L.
  Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
  C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
  Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
  A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
  ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
  A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
  A:Reference number: A86141; MUID:21016719
A:Accession: C96673
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: GB:AE005173; NID:g8099780; PIDN:AAD38271.2; GSPDB:GN00141
C:Genetics:
A:Gene: F13011.27
A:Map position: 1

```

Query Match

Best Local Similarity 27.6%; Score 177; DB 2; Length 348;
Matches 54; Conservative 32; Mismatches 62; Indels 48; Gaps 8;

```

Qy 223 NQISWLMQKV-----DSE---DDKFGQ---RFLDSSQYKFNILRYRVFGP 263
      : : || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 NETSGIMEIIGHDMHGFYDPDSSVOLSDSGHKEAQIRMEES-----LRPAGV--- 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 264 GYSTGLETTEKFEVSLDLKPGKVIYDVGCGIGGDFYMAENDVEVGIIDLSIMISF 323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 -----TDEEEKRI-----KKVYDVGCGIGGSSRYLASKFGAFCIGITLSPVQAKR 164
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 324 A--LERSIGLKAVEFEVADCTKKDYSPNSFDVYSRDTILHIDKRALPFRSFKWLKPG 381
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 ANDLAAQSLAHKASFYVAADALDPFEDGKFDLVMSGESGHEHMDKAKFYKELVRAVAPG 224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 382 GKVLISDYCK--SAG 394
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 GRITVTWCHRNLSAG 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

S18533

eryG protein - Saccharopolyspora erythraea

C:Species: Saccharopolyspora erythraea

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 22-Oct-1999

C:Accession: S18533; S16747

R:Haydock, S.F.; Dowson, J.A.; Dillon, N.; Roberts, G.A.; Cortes, J.; Leadlay, P.F.

Mol. Gen. Genet. 230, 120-128, 1991

A:Title: Cloning and sequence analysis of genes involved in erythromycin biosynthesis in

methylintransferases.

A:Reference number: S18530; MUID:92079886

A:Accession: S18533

A:Molecule type: DNA

```

A:Residues: 1-306 <HAY>
A:Cross-references: EMBL:X60379; NID:g48941; PIDN:CAA42929.1; PID:g581650
C:Genetics:
A:Gene: eryG
A:Start codon: GTG
C:Superfamily: b10c homology
F:82-187/Domain: b10c homology <BIOC>

```

Query Match

Best Local Similarity 27.8%; Score 176.5; DB 2; Length 306;
Matches 60; Conservative 36; Mismatches 81; Indels 39; Gaps 9;

```

Qy 264 GYSTG--GL-ETTKFEVSKL-----DLKPGKVIYDVGCGIGGDFYMAE--NYDVEVVGID 315
      : : || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 53 GYMKPGCAGLEANGELANGLAEAGISEGDEVLDVGFGLGADQFFMLETKPARIYGV 112
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 316 LSIMISFALERS--IGLKAVEFEVADCTKKDYSPNSFDVYSRDTILHIDKRALPFRS 373
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 LTPSHVRIASERARERENVQDRLOPKFEGSATDLPFGAETFDVTSLESAHLYEPRTDFEKG 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 374 FHKMLKPGKVLISDYCK-----SAGTPSAEFAAIIYRGYDLH-----DYKAY 417
      : : |||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 AFEVLKPGVLAIGDIIPLDLREPSSGDPK-----LAPORSGSLSGIPEVKNVPRETY 227
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 418 GKMLKDAFVEV-----IAENRTDQFIQVLQKE 445
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 AKQAREAGFVDEYKSVNDNMPEWMLYLRKLQDE 263
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: September 20, 2001, 16:20:09
Job time: 167 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:19:23 ; Search time 14.71 Seconds

(without alignments)
1150.388 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598
Sequence: 1 MAASAMGYLOEREFVKKYWI.....KLVRTGEQOGLFIARKM 494

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	176	6.8	381	1	CFA_ECOLI
2	170	6.5	382	1	ERCG_YEAST
3	161.5	6.2	378	1	ERCG_SCHPO
4	148.5	5.7	376	1	ERCG_CANAL
5	144	5.5	244	1	YXBB_BACSU
6	144	5.5	251	1	UBIE_ECOLI
7	135.5	5.2	345	1	YTJ7_STRFR
8	135	5.2	234	1	MENH_BACST
9	134.5	5.2	232	1	GRCH_LACLA
10	128	4.9	240	1	UBIG_ECOLI
11	128	4.9	417	1	YAT1_SYNP6
12	127.5	4.9	209	1	UBIG_SALTY
13	126	4.8	212	1	Y829_SYNY3
14	123.5	4.8	197	1	TEHB_ECOLI
15	123	4.7	231	1	Y095_HAEIN
16	120	4.6	254	1	Y912_HAEIN
17	119	4.6	1021	1	DPOH_NEUCR
18	118.5	4.6	247	1	Y0EM_BACSU
19	117.5	4.5	230	1	P1MT_PYRHO
20	117.5	4.5	394	1	YLPD_PSEBP
21	117	4.5	255	1	BIOC_SEPMA
22	116.5	4.5	286	1	CEAL_SEPMA
23	115	4.4	287	1	CEAL_MYCTU
24	115	4.4	302	1	CEFA2_MYCTU
25	114.5	4.4	1091	1	CIC2_HUMAN
26	113.5	4.4	1106	1	CIC2_RABIT
27	112	4.3	306	1	Y028_RICPR
28	111	4.3	3058	1	POLG_PEMVC
29	110.5	4.3	1435	1	YPX3_MYCPU
30	110	4.2	243	1	YX42_MYCTU
31	110	4.2	1091	1	CIC2_RAT
32	109.5	4.2	626	1	HTPG_BACSU
33	109.5	4.2	699	1	HS81_ARATH

34	106.5	4.1	251	1	BIOC_ERWHE
35	106.5	4.1	1675	1	POL_RTBVP
36	106	4.1	261	1	TAM_MYCTU
37	105	4.1	621	1	HTPG_HELPY
38	105.5	4.1	621	1	HTPG_HELPY
39	105	4.1	715	1	Y2P2_CANFA
40	105	4.0	1071	1	VATA_YEAST
41	104.5	4.0	203	1	PMTA_RHOSH
42	104.5	4.0	366	1	Y906_METJA
43	104.5	4.0	1294	1	RRPO_MCMVM
44	103.5	4.0	1700	1	HS81_ARATH
45	103.5	4.0	1376	1	RPOB_RICTY

ALIGNMENTS

RESULT 1

ID	CFA_ECOLI	STANDARD	PRT	381 AA.
AC	P30010			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE (EC 2.1.1.79)			
DE	(CYCLOPROPANE FATTY ACID SYNTHASE) (CFA SYNTHASE).			
GN	CFA OR CDEA.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A., SEQUENCE OF 1-8, AND CHARACTERIZATION.			
RC	SPRATN-K12			
RC	MEDLINE=93075691; PubMed=1445840;			
RA	Wang A.-Y., Grogan D.W., Cronan J.E. Jr.,			
RT	"Cyclopropane fatty acid synthase of Escherichia coli: deduced amino acid sequence, purification, and studies of the enzyme active site."			
RL	Biochemistry 31:11020-11028(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPRATN-K12 / RR28;			
RA	Eberhardt S.M.R., Richter G., Gimbel W., Werner T., Bacher A.;			
RL	Submitted (NOV-1992) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	SPRATN-K12 / MG1655;			
RC	MEDLINE=97426617; PubMed=9278503;			
RA	Bialtner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12."			
RL	Science 277:1453-1474(1997).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	SPRATN-K12;			
RC	MEDLINE=97251357; PubMed=9097039;			
RA	Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,			
RA	Itoh T., Kasai K., Kashimoto K., Kimura S., Kitahawa M.,			
RA	Kitahawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,			
RA	Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,			
RA	Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,			
RA	Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,			
RA	Yamamoto Y., Horiuchi T.;			
RT	"A 570-Kb DNA sequence of the Escherichia coli K-12 genome			
RT	corresponding to the 28.0-40.1 min region on the linkage map."			
RL	DNA Res. 3:363-377(1996).			
CC	-1- FUNCTION: TRANSFERS A METHYLENE GROUP FROM S-ADENOSYL-L-METHIONINE			
CC	TO THE CIS DOUBLE BOND OF AN UNSATURATED FATTY ACID CHAIN			
CC	RESULTING IN THE REPLACEMENT OF THE DOUBLE BOND WITH A METHYLENE			
CC	BRIDGE.			
CC	-1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + PHOSPHOLIPID			

```
CC OLEFINIC FATTY ACID - S-ADENOSYL-L-HOMOCYSTEINE + PHOSPHOLIPID
CC CYCLOPROPANE FATTY ACID.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M98330; AAA23562.1; -
DR EMBL: X69109; -; NOT_ANNOTATED_CDS.
DR EMBL: AE000261; AAC74733.1; -
DR EMBL: D90809; BAA15428.1; -
DR EMBL: D90810; BAA15437.1; -
DR PIR: A44292; A44292.
DR EcoGene: EGI1531; cfa.
DR Pfam: PF00972; Flavi_N5; 1.
KM Transferase: Methyltransferase; Lipid synthesis.
FT INIT_MET 0 0
FT CONFLICT 1 1 S -> R (IN REF. 2).
FT CONFLICT 7 7 E -> G (IN REF. 2).
FT CONFLICT 24 24 S -> N (IN REF. 2).
FT CONFLICT 38 38 I -> T (IN REF. 2).
SQ SEQUENCE 381 AA; 43777 MW; 1F07B220C7E08ADF CRC64;

Query Match 6.88; Score 176; DB 1; Length 381;
Best Local Similarity 29.38; Pred. No. 1,2e-05;
Matches 65; Conservative 27; Mismatches 80; Indels 50; Gaps 12;

QY 184 FYRKIF-----ECHMODSGNSYELSLIGCKIGAVYKSKKNQISMLQKVSSED 236
DB 74 FFSKVLKAGLENDLPNFKD-----TLRTAGARLEN-LQSKR---AWYK--EHY 119
QY 237 DKG---FQRLDSSQYKFNILRYERFGPGYSTGLETT-----KEFYSKIDLPKQK 288
DB 120 DLGNDLFSLMD-----PFQYSCAY---WMDADNLESAQOKLMICEKILQKFGMR 169
QY 289 VLDVGGCIGGDFYMAENYVEVVGIDLSIMISFALERISIGLKAVEFEVADCTKKDYP 348
DB 170 VLDIGCGGGLAHYMASNYVSVGVISAEQKMADEKGEGLDVTLLD-----DYR 222
QY 349 E--NSFDVYSRDTILHIQDK--PALFRSFHKMLKPGKVL 386
DB 223 DLNDQFDRIYVGMFEHVGPKNDYFAVYDRLNKPBGIFLL 264

RESULT 2
ERG6_YEAST STANDARD; PRT; 382 AA.
AC P25087:
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE DELTA(24)-STEROL C-METHYLTRANSFERASE (EC 2.1.1.41).
GN ERG6 OR SED6 OR ISEL OR LISI OR YML08C OR YW9571.10C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetales; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94262330; PubMed=8203167;
RA Hardwick K.G., Pelham H.R.B.;
RT "Se6 is identical to ERG6, and encodes a putative methyltransferase
RL required for ergosterol synthesis.";
RL Yeast 10:265-269(1994).
RP SEQUENCE FROM N.A.
RX MEDLINE=94312403; PubMed=8038180;
```

```
RA Melinda A.A., Beavis A.D., Trumbly R.J.;
RT "Mutations in LISI (ERG6) gene confer increased sodium and lithium
RT uptake in Saccharomyces cerevisiae.";
RL Biochim. Biophys. Acta 1193:107-117(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RC Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-257 FROM N.A.
RX MEDLINE=91285426; PubMed=2060792;
RA Hussain M., Leonard J.;
RT "Characterization of PDR4, a Saccharomyces cerevisiae gene that
RT confers pleiotropic drug resistance in high-copy number: identity
RT with YAP1, encoding a transcriptional activator.";
RL Gene 101:149-152(1991).
RN [5]
RP ACETYLATION.
RA Garets J.I., Futcher B., Kobayashi R., Latter G.I., Schwender B.,
RA Volpe T., Warner J.R., McLaughlin G.S.;
RL Submitted (SEP-1994) to the SWISS-PROT data bank.
CC -1- FUNCTION: METHYLTRANSFERASE REQUIRED FOR ERGOSTEROL SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 5-ALPHA-CHOLESTA-
CC 8,24-DIEN-3-BETA-OL -> S-ADENOSYL-L-HOMOCYSTEINE + 24-METHYLENE-
CC 5-ALPHA-CHOLEST-8-EN-3-BETA-OL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74249; CA52308.1; -
DR EMBL: S72460; AAB31378.1; -
DR EMBL: Z49810; CAA89944.1; -
DR EMBL: X53830; CAA37826.1; -
DR PIR: PEO402; PEO402.
DR PIR: S17001; S17001.
DR PIR: S35982; S35982.
DR PIR: S42003; S42003.
DR SWISS-2DPAGE; P25087; YEAST.
DR YEPD; 6410; -
DR SGD: S0004467; ERG6.
KM Sterol biosynthesis; Transferase; Methyltransferase; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT CONFLICT 379 379 E -> EE (IN REF. 2).
SQ SEQUENCE 382 AA; 43299 MW; D76BAE9E3D9CD71B CRC64;

Query Match 5.58; Score 170; DB 1; Length 382;
Best Local Similarity 23.68; Pred. No. 3,2e-05;
Matches 65; Conservative 44; Mismatches 127; Indels 40; Gaps 9;

QY 190 KECHMODSGNSYELSLIGCKIGAVYKSKKNQ--ISMLQKVSSEDDKG----- 239
DB 14 RELH-GDGIKTKGLSL-----MSKNNSAQKAVQKYLBNMDGFRDKDAEERLE 63
QY 240 -----FQRLDSSQYKFNILRYERFGPGYSTGLETTKEFVS-KIDLKPGQKY 289
DB 64 DYNEATHSYNNVTIDREYGMGSSFFHSR-FYKGESEFASIAHREHYLAKGIQKGDIV 122
QY 290 LDVGGCIGGDFYMAENYVEVVGIDLSIMISFA--LERSIGLKCAVEFEVADCTKKOY 347
DB 123 LDVCGGCGGPARIEARTGCVGLNNNDYQIAKAYAKKYVLSQOMDVKGDPMKND 182
QY 348 PENSFDVYSRDTILHIQDKPALFRSFHKMLKPGKVLSDYCKS-----AGTSAEFAV 403
DB 183 EENTFDKVVALEATCAKPKLEGVSELYKVLKPGGTFAVYEWMTDKYDENNEHKKIAY 242
```

OY 404 IROGC---YDLHDYKANGKMLKADGFEVIAENRTD 436
 Db 243 EIELGDGIPKMFHVDVARKALKNGCFEVLASEDLAD 278

RESULT 3

ERG6_SCHPO STANDARD; PRT; 378 AA.
 ID ERG6_SCHPO
 AC 014321: P78782;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PROBABLE DELTA(24)-STEROL C-METHYLTRANSFERASE (EC 2.1.1.41).
 GN SPBC16E9.05.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Volckert G., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 55-378 FROM N.A.
 RC STRAIN=PR745;
 RX MEDLINE=98162722; PubMed=9501991;
 RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;
 RT "Identification of open reading frames in Schizosaccharomyces pombe
 RT cDNAs.";
 RN DNA Res. 4:363-369(1997).
 CC -1- FUNCTION: METHYLTRANSFERASE REQUIRED FOR ERGOSTEROL SYNTHESIS (BY
 CC SMITHIARITY).
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 5-ALPHA-CHOLESTA-
 CC 8,24-DIEN-3-BETA-OL = S-ADENOSYL-L-HOMOCYSTEINE + 24-METHYLENE-
 CC 5-ALPHA-CHOLEST-8-EN-3-BETA-OL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z69759; CAB16897.1; -;
 DR EMBL: D89131; BA13793.1; -;
 DR InterPro: IPR000339; -;
 DR Pfam: PF01209; Ublc_methyltran; 1.
 KW Sterol biosynthesis; Transferase; Methyltransferase.
 FT CONFLICT 55 NHENESRD -> SYRCDPLES (IN REF. 2).
 SQ SEQUENCE 378 AA: 42867 MW: FA4D3D82A1CE03D6 CRC64;

Query Match 6.2%; Score 161.5; DB 1; Length 378;
 Best Local Similarity 21.9%; Pred. No. 0.00013;
 Matches 66; Conservative 62; Mismatches 114; Indels 59; Gaps 15;

OY 217 KSKRNQNOISMLWOK---VDSDDK---GFQRFIDS-----SOYKENSILRYERV 260
 Db 41 OSRLQGEFFEF-WDRNHNENESEEDRRARIDGYKSVNSYDLATDLYEYGSQSFHFSR- 98
 OY 261 FGPYVSTGGLTKEFVS-KLIDLPGOKVLVDCGIGCGFYMAENDVAVGIDLSIN 319
 Db 99 FYKGEAFQAQSIARHHEHYLAAYMGIKPGSRVLVDCGVGVPARETTEFTGCMLVGNNDY 158
 OY 320 MIS-----FALERSIGLKAVEFEVADCTKKDYPENSPDVITSRDTLHIODKRALFRSFH 375
 Db 159 QISCNNAVAVARNLDKKGV-FVKGDFMHMPFEDNTEFYVAIAETVAPARPLEGYSIEIF 216
 OY 376 KWLPGKG-----VLISDYKKSAGTSAEFAAYIRQGYLDHDKVAYGKMLKADGFEV 429

Db 217 RVLPGGVGYEYEWMSDDY--DSSIPKH-----REIAYNIEGDIPIQVYRRCDAVEA 268
 OY 430 IAENRDQFIQVLOKELDALBOEKDDFIDPSEEDYNDIVGWMKALY-RTTGEEOQGL 488
 Db 269 IK-----KVGFNLEE-----DILTIDHNDPL--PWYIPLTGLITKCONIMDV 309

OY 489 F 489
 Db 310 F 310

RESULT 4

ERG6_CANAL STANDARD; PRT; 376 AA.
 ID ERG6_CANAL
 AC 074198;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DELTA(24)-STEROL C-METHYLTRANSFERASE (EC 2.1.1.41).
 GN ERG6.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Microsporica; Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98253976; PubMed=9593144;
 RX Jensen-Petgakes K.L., Kennedy M.A., Lees N.D., Barbuch R., Koegel C.,
 RA Bard M.;
 RT "Sequencing, disruption, and characterization of the Candida albicans
 RT sterol methyltransferase (ERG6) gene: drug susceptibility studies in
 RT erg6 mutants.";
 RN Antimicrob. Agents Chemother. 42:1160-1167(1998).
 CC -1- FUNCTION: METHYLTRANSFERASE REQUIRED FOR ERGOSTEROL SYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 5-ALPHA-CHOLESTA-
 CC 8,24-DIEN-3-BETA-OL = S-ADENOSYL-L-HOMOCYSTEINE + 24-METHYLENE-
 CC 5-ALPHA-CHOLEST-8-EN-3-BETA-OL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF031941; AAC26626.1; -;
 KW Sterol biosynthesis; Transferase; Methyltransferase.
 SQ SEQUENCE 376 AA: 43085 MW: 369D094AFCDDED0A0 CRC64;

Query Match 5.7%; Score 148.5; DB 1; Length 376;
 Best Local Similarity 24.0%; Pred. No. 0.0011;
 Matches 63; Conservative 42; Mismatches 102; Indels 55; Gaps 12;

OY 229 WQKVDSEDDKGFQFLDSSO-----YKFSILRYERFEGPVGSTGLETT 274
 Db 53 WDGGSKNDEE-KRLNDYSQTLNHYNLVTPFEYGMSSSFHSRY-KGAFAQATARRH 110
 OY 275 KEFVS-KLIDLPGOKVLVDCGIGCGDFYMAENDVAVGIDLSINMISFALERS----- 328
 Db 111 EHLFAHKNNLNNKMLVLDVCGVGCGPGRRETRFDCEIVG-----LNNNDYQIERANHYAK 166
 OY 329 -IGLKCAVEFEVADCTKKDYPENSPDVITSRDTLHIODKRALFRSPFKMLKPGKYLIS 387
 Db 167 KYHLHDKLSYVKGFMQMDPEPESFDAYALEAVHVLVEGVSELYKVLKPGG----- 221
 OY 388 DYKKSAGTPSAEFAAY--IROGYD---LHDVAYGKMLKADGFEVIAENRPTDQIOV 441
 Db 222 -----IFGVLEVWMTDKDIDETNEHRKATAYGLEVGD-GIPKYSKVAEQALKN 269
 OY 442 LQKELDLEQEKDFIDPSEE 463

Db 270 VGFE---LEYOKD--LADVDE 286

RESULT 5

ID YXBB_BACSU STANDARD; PRT; 244 AA.

AC P46326;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE HYPOTHEICAL 28.2 KDA PROTEIN IN ALDX-ASNH INTERGENIC REGION.

GN YXBB OR VE/CR.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / BGSC1A1;

RX MEDLINE=96093926; PubMed=7584049;

RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;

RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis genome between the gnt and lol operons.";

RL DNA Res. 2:61-69(1993).

CC -1- SIMILARITY: TO E.COLI YAFE.

CC -1- SIMILARITY: TO METHYLTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AB00554.1; BAA21596.1; -

DR EMBL: Z99124; CAB16025.1; -

DR Subtilist; BG11352; yxsb.

DR Interpro: IPR000339; -

DR Pfam: PF01209; Uble_methyltran; 1.

KW Hypothetical protein.

SO SEQUENCE 244 AA; 28222 MW; 5B008371AFD3D02B CRC64;

SO

Query Match 5.5%; Score 144; DB 1; Length 244;

Best Local Similarity 20.4%; Pred. No. 0.0013;

Matches 55; Conservative 49; Mismatches 118; Indels 48; Gaps 8;

QY 232 VSEDDKGFQFPLDSSQKFNSTL--RYERVFQGVSTGGLTTKEFVSKLDKPGQKV 289

DB 2 ITAECHRTYEFVLYOSLYPLGLARLSRMSP-----KMKVI 40

QY 290 LDVGGGIGGDDFYAENVYDVEVGDIDISIMISFALF--RSIGLKAVEFEVADCTKKDY 347

DB 41 IDMGTEGGYSLQAKTNAHVAVDINPMHETAQEAKKSGVSSLSIDLEDVHHLST 100

QY 348 PENSFDVYSRDTLTIQDPALEFRSFHMKLPGKVLISDYCSACT-----PSA 398

DB 101 ADQYADRTVYSCLHMHEDVVKGLKECYRVLAPOGKLVILDTFPMQSHLEIMKQIKEP 160

QY 399 EFAAYIR--QRGYDLHVAAYGKMLKDAQFVEVIAEN--RFDQFQVLOKELDALEQ 452

DB 161 EYRFVAREAEESYSFEDIQF--VODAGIPNYSLETFFHFLPEDFTESLDELADAPLWE 217

QY 453 KDDFIDFSEEDYNDIVDGKAKLVRTTEG 482

DB 218 QNDQSTDHETESVT-----MLTIKEKEG 242

Db

RESULT 6

UBIE_ECOLI STANDARD; PRT; 251 AA.

AC P27851;

DT 01-AUG-1992 (Rel. 23, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE UBIQUINONE/MENAOQUINONE BIOSYNTHESIS METHYLTRANSFERASE UBIE

DE (EC 2.1.1.-).

GN UBIE.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=92358234; PubMed=1379743;

RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;

RT "Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 to 86.5 minutes.";

RL Science 257:771-778(1992).

RL [2]

RP REVISION TO 101.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP CHARACTERIZATION, AND VARIANT UBIE401.

RX MEDLINE=97197541; PubMed=9045837;

RA Lee P.T., Hsu A.Y., Ha H.T., Clarke C.F.;

RT "A C-methyltransferase involved in both ubiquinone and menaquinone biosynthesis: isolation and identification of the Escherichia coli ubie gene.";

RL J. Bacteriol. 179:1748-1754(1997).

CC -1- FUNCTION: CONVERTS DMOH2 INTO DMOH2 AND DMKH2 INTO MKH2.

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 2-POLYPRENYL-6-METHOXY-1,4-BENZOQUINOL = S-ADENOSYL-L-HOMOCYSTEINE + 2-POLYPRENYL-3-METHYL-6-METHOXY-1,4-BENZOQUINOL.

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DIMETHYLMENAQUINOL = S-ADENOSYL-L-HOMOCYSTEINE + MENAQUINOL.

CC -1- PATHWAY: INVOLVED IN BOTH UBIQUINONE AND MENAQUINONE BIOSYNTHESIS.

CC -1- SIMILARITY: BELONGS TO THE UBIE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M87049; AAA67628.1; -

DR EMBL: AE000459; AAC76636.1; -

DR PIR: S30722; S30722.

DR EcoGene: EG11473; ubie.

DR Interpro: IPR000339; -

DR Pfam: PF01209; Uble_methyltran; 1.

DR PROSITE: PS01183; UBIE_1; 1.

DR PROSITE: PS01184; UBIE_2; 1.

KW Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase; Methyltransferase.

FT VARIANT 142 142 G -> D (IN UBIE401 ALLELE, DEFECTIVE).

FT CONFLICT 101 101 P -> L (IN REF. 1).

SO SEQUENCE 251 AA; 28057 MW; 646F3E09A0075A9E CRC64;

SO

Query Match 5.5%; Score 144; DB 1; Length 251;

Best Local Similarity 26.8%; Pred. No. 0.0014;

Matches 52; Conservative 36; Mismatches 88; Indels 18; Gaps 7;

QY 283 LKPGQKVLVGGIGG--GDFYMAENVYDVEVGDIDISIMISFALF--RSIGLKAVEFE 338


```

Db      61 VRRCQTVLADLAGTGLTAFAFSRLVGETGKRVYLVADINESPMKRGKRLRNIGVANEYV 120
      339 VADCTKDYFENSPFVYISRDITLHIDKPALFSPFKMLKPGKVLISYCSAGTSPA 398
      121 QANNEALPFDPNTDCTTISFGLRNVTDKOKALRSKRYLKGGRLLVLEFSKIIIEPLS 180
      399 EFAAYIRQGYDLHDVAKYKML-KDAGFEVIAENRTDQFIQVLOKELDALEQEKDQFI 457
      181 K-AV---DAVSFHLPRISLVANADSYRYLAESIR-----MHPDQDTLKAMQD-- 227
      458 DDFSEEDYNDIVDG 471
      228 AGESEVDYVNLTAG 241

```

```

RESULT 7
YT37_STRFR STANDARD; PRT; 345 AA.
ID YT37_STRFR
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE HYPOHETICAL 37.1 KDA PROTEIN IN TRANSPOSON TNA556.
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-TNA556;
RX MEDLINE-90185236; PubMed-2155856;
RA Siemieniak D.R., Slightom J.L., Chung S.T.;
RT TNA556: a class-II transposon related to Tn3.
RL Gene 86.1-9(1990).
CC -I- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M29297; AAA88564.1; -.
CC PIR: J00429; J00429.
CC InterPro: IPR000339; -.
CC Pfam: PF01209; Uble_methyltran; 1.
CC Hypothetical protein; Transferase; Methyltransferase;
CC Transposable element.
CC KW
CC SEQUENCE 345 AA; 37113 MW; 027CDA5B3F8AA7CC CRC64;

```

```

Query Match 5.2%; Score 135.5; DB 1; Length 345;
Best Local Similarity 26.9%; Pred. No. 0.009;
Matches 45; Conservative 22; Mismatches 67; Indels 33; Gaps 6;
QY 281 LDLPGRKVLVDGCGIGG--GDFYMAENVYVEVGIDLSIMISFALERSGLCAVEFE 338
DB 122 LDAPRGESALDLGGCGPGTDLGTAKAVSPSGRVIGIDSSQEMVQARRRENLP-AVEVE 180
QY 339 VADCTKDYFENSPFVYISRDITLHIDKPALFSPFKMLKPGKVLISYCSAGTSPA 389
DB 181 LGDHTHTLPLEGGSIDCARTDQVLOHVADPQAALAEARRVLRPGSLRVWGEPMDSLTIDY 240
QY 390 CKSAGTSPSAEFAAYIRQGYDLHDV-----KAYGKMLKDAGF 426
DB 241 -----PDLEVS-----RAYTRHVTDKIVRNGVIGRQLARLALDAGF 276

```

RESULT 8

```

MENH_BACST STANDARD; PRT; 234 AA.
ID MENH_BACST
AC 086169;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2-HEPTAPRENYL-1,4-NAPHTHOQUINONE METHYLTRANSFERASE (EC 2.1.1.-)
DE (MENAQUINONE BIOSYNTHESIS METHYLTRANSFERASE).
GN MENH OR MENH.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN-ATCC 10149;
RX MEDLINE-97284706; PubMed-9139683;
RA Koike-Takeshita A., Koyama T., Ogura K.;
RT "Identification of a novel gene cluster participating in menaquinone
RT (vitamin K2) biosynthesis. Cloning and sequence determination of the
RT 2-heptaprenyl-1,4-naphthoquinone methyltransferase gene of Bacillus
RT stearothermophilus."
RL J. Biol. Chem. 272:12380-12383(1997).
CC -I- FUNCTION: CONVERTS DMK2 INTO MKH2.
CC -I- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DEMETHYLMENAQUINOL =
CC S-ADENOSYL-L-HOMOCYSTEINE + MENAQUINOL.
CC -I- PATHWAY: LAST STEP IN MENAQUINONE BIOSYNTHESIS.
CC -I- SIMILARITY: BELONGS TO THE UBIE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D87054; BAA32500.1; -.
CC InterPro: IPR000339; -.
CC Pfam: PF01209; Uble_methyltran; 1.
CC PROSITE: PS01184; UBIE_2; FALSE_NEG.
CC Menaquinone biosynthesis; Transferase; Methyltransferase.
CC KW
CC SEQUENCE 234 AA; 27128 MW; 8C1CDA06A525F953 CRC64;

```

```

Query Match 5.2%; Score 135; DB 1; Length 234;
Best Local Similarity 21.9%; Pred. No. 0.0057;
Matches 47; Conservative 42; Mismatches 80; Indels 46; Gaps 7;
QY 246 SSOY-KFNSILRYERFVGPVSTGLETTEKFEYSKDLKPGQKVLVDGCGIGGDFYMA 304
DB 17 SAHIDRNVSVISFR-----HLKWRKDVRRNRNVQKKGKALDVCCTGTDWITIALA 66
QY 305 ENIDVE--VVGIDLSIMISFALF--RSIGLKAVEFVADCTKDYFENSPFVYISRDIT 360
DB 67 EAVGPEKRYVGLPSEMMKLKGVGEQKVARGLH-NVKLIHGNAMOLPPDPSFDVYITGFG 125
QY 361 ITHIDKPALFSPFKMLKPGKVLISDYCSAGT- 396
DB 126 LRNPVDYMTVLKEMHKRYTKRGITV-----CLETSQPLFLFGFRQLYYTFRFPIMPLFKLL 181
QY 397 -SAEFAAYIRQGYDLHDVAKYKMLKDAGFVEV 429
DB 182 AKSYEEYSWLGESARERPPGRDELAEFRAGVFVDV 216

```

```

RESULT 9
GRCB_LACLA STANDARD; PRT; 252 AA.
ID GRCB_LACLA
AC P49016;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```

DT 30-MAY-2000 (Rel. 39, last annotation update)
 DE PROBABLE MENAQUINONE BIOSYNTHESIS METHYLTRANSFERASE (EC 2.1.1.-)
 OS (GERC2 PROTEIN HOMOLOG)
 OC Lactococcus lactis (subsp. lactis) (Streptococcus lactis)
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CC2;
 RX MEDLINE=93374846; PubMed=8366036;
 RA Geller B.L., Ivey R.G., Trempey J.E., Hettlinger-Smith B.;
 RT Cloning of a chromosomal gene required for phage infection of
 RT Lactococcus lactis subsp. lactis CC2.
 RL J. Bacteriol. 175:5510-5519(1993).
 CC -1- FUNCTION: CONVERTS DMK2 INTO MKH2.
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DEMETHYLMENAQUINOL =
 CC S-ADENOSYL-L-HOMOCYSTEINE + MENAQUINOL.
 CC -1- PATHWAY: LAST STEP IN MENAQUINONE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE UBIE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L14679; AAA03166.1; -
 DR InterPro: IPR000339; -
 DR Pfam: PF01209; Ubie_methyltran; 1.
 DR PROSITE: PS01183; UBIE.1; 1.
 DR PROSITE: PS01184; UBIE.2; 1.
 KM Menaquinone biosynthesis; Transferase; Methyltransferase.
 SQ SEQUENCE 252 AA; 28449 MW; 7D971E3A4BD804ED CRC64;

Query Match 5.2%; Score 134.5; DB 1; Length 252;
 Best Local Similarity 21.5%; Pred. No. 0.0069;
 Matches 52; Conservative 44; Mismatches 97; Indels 49; Gaps 7;

QY 215 YVNS--KKNOQISLWQKXDESDKGFQRFLLSSQYKFNLSILRYEFGPGYSTGGL 272
 | | | | | : : : : : | : : : : : | : : : : : | : : : : : |
 | | | | | : : : : : | : : : : : | : : : : : | : : : : : |
 DB 3 YNKSMTVNERVEQELFNSISDSD-----KMAAISFKO-----HDLW 41

QY 273 TTKEFVSKLKGQKVLVDGCGGCGGFYMAENY--DVEVGDLSINMISFALER-SI 329
 | | | | | : : : : : | : : : : : | : : : : : | : : : : : |
 | | | | | : : : : : | : : : : : | : : : : : | : : : : : |
 DB 42 RAKTKRMGDL-TGLSLDLCGGTGDWTFDLSVSGVKGVLGDFSENMELAKAKLKE 100

QY 330 GLKCAVEFEVADCTKKDYPENSPDYISRDITLHIDKPALFRSFHMLKPGKVL---- 385
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 DB 101 EAKKNIEFLGNMMLPFERKSGFYVITGYGLRTPYLVTKLFIKFLVLRPGGVVCIET 160

QY 386 -----ISDYCKSAGTSPSAEFPAVYTRQGYDLHDYKAYGKMKLDAGEV 427
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 DB 161 SHPTLPYKQAFELYKRNVMPLFGKVFPAKSLKEYQWLQKSAEDFPDPAKLEELFRKAGFV 220

QY 428 EV 429
 |
 DB 221 AV 222

RESULT 10
 UBIG_ECOLI STANDARD; PRT; 240 AA.
 AC P17993; P76924;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, last sequence update)
 DT 01-OCT-2000 (Rel. 40, last annotation update)
 DE 3-DEMETHYLBIOQUINONE-9-3-METHYLTRANSFERASE (EC 2.1.1.64)
 DE (3,4-DIHYDROXY-5-HEXAPRENYLBENZOATE METHYLTRANSFERASE) (DHMB
 METHYLTRANSFERASE).

GN UBIG OR PUPX.
 OS Escherichia coli.
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93123968; PubMed=1479344;
 RA Wu G., Williams H.D., Zamanian M., Gibson F., Poole R.K.;
 RT Isolation and characterization of Escherichia coli mutants affected
 RT in aerobic respiration: the cloning and nucleotide sequence of ubig.
 RT Identification of an S-adenosylmethionine-binding motif in protein,
 RT RNA, and small-molecule methyltransferases.
 RL J. Gen. Microbiol. 138:2101-2112(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OV6;
 RX MEDLINE=88201664; PubMed=2834621;
 RA Hussain K., Elliott E.J., Salmood G.P.C.;
 RT "The para- mutant of Escherichia coli also carries a gyrAa mutation.
 RT Mol. Microbiol. 1:259-273(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T.,
 RA Ikemoto K., Inada T., Isono K., Itoh T., Kanai K., Kasai H.,
 RA Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
 RA Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
 RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,
 RA Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE +
 CC 3-DEMETHYLBIOQUINONE-9 = S-ADENOSYL-L-HOMOCYSTEINE +
 CC BIOQUINONE-9.
 CC -1- PATHWAY: BIOQUINONE BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00544; CA68610.1; -
 DR EMBL: M87509; AAA24714.1; -
 DR EMBL: AE000313; AAC75292.1; -
 DR EMBL: D90854; BAA16049.1; -
 DR EMBL: D90855; BAA16051.1; -
 DR PIR: S03757; S03757.
 DR PIR: A47682; A47682.
 DR EcGene: EG1143; ubig.
 DR InterPro: IPR000339; -
 DR Pfam: PF01209; Ubie_methyltran; 1.
 KM Ubiquinone biosynthesis; Transferase; Methyltransferase.
 SQ SEQUENCE 240 AA; 26555 MW; D4EB4707A06F5613 CRC64;

Query Match 4.9%; Score 128; DB 1; Length 240;

Best Local Similarity 24.8%; Pred. No. 0.019;
Matches 59; Conservative 38; Mismatches 75; Indels 66; Gaps 13;

```

OY 217 KSKKNONQISLMQKVDSEDKGFQ---RFLDSSQYKFNILYERFEGP-----263
DB 5 KSPVNHN-----VHEETAKFEAVASRWMD-----LEGEFKLHINIPRL 45
OY 264 GYVS--TGLETTKEFVSKDLKPGOKVLDVCGGIGGDFYMAENVYD---VVGIDLSI 318
DB 46 GYIERRAGGLE-----GKRVLDVGC---GGG--ILASSMAREGATVAGLDMGF 88
OY 319 NMISFALERSIGLKCAVEFEVADCTKKDYP---ENSEVDIYSRDTILHIODKPALFRSFH 375
DB 89 EPILOYA--KLHALSFGIOVDVYQETVEEHAHAKHAGQYDVYTCMELHEVDPQSVVRACA 146
OY 376 KWLKPGKVLISDYCKSA---GTPSAEFAVYIRORGYDLDHVKAYGKMKDKAGFEV 428
DB 147 QLVKPGGVFFSTLNRNGKSWLMAVGAELYILRMVPG--THDVKKFKIPALLGWD 202

```

```

RESULT 11
YATL_SYNP6
ID YATL_SYNP6 STANDARD; PRT; 417 AA.
AC P08442;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)
DE ATP SYNTHASE SUBUNITS REGION ORF 1.
OS Synecococcus sp. (strain PCC 6301). (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87311713; PubMed=3041005;
RA Cozens A.L., Walker J.E.;
RT "The organization and sequence of the genes for ATP synthase subunits
in the cyanobacterium Synecococcus 6301. Support for an
endosymbiotic origin of Chloroplasts."
RL J. Mol. Biol. 194;359-383(1987).
CC -1- SIMILARITY: TO M.TUBERCULOSIS RV2067C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X05302; CAA28921.1; -.
DR PIR; S07286;
KM Hypothetical protein.
SQ SEQUENCE 417 AA; 47531 MW; 4F96D4F5E7C0F99B CRC64;

```

```

Query Match 4.9%; Score 128; DB 1; Length 417;
Best Local Similarity 25.3%; Pred. No. 0.041;
Matches 55; Conservative 29; Mismatches 87; Indels 46; Gaps 8;
OY 288 KVLVDVGGGIGGDFYMAE--NIDVEVVGIDLSINMISFALER--SIGLKCAVEFEVADCTK 344
DB 74 RILDAGGCTGVSTDYLAHLNPSAETIADISAGTLAQAQECQSRGVAADRTHFOQLSYD 133
OY 345 KDYPENSFDVYISRDTILHIODKPALFRSFHKKLPGKVLISDYCKSAGPSAEFAVYI 404
DB 134 VAQIPGFEQDQNCYGVHLHLEDPDRGLAALASKLAPGILHIIFYAELG-----RAEI 186
OY 405 RQ-----RGYDLHDVKAYGKMKDKAGVEVIAENFTDOFIQVLRKDALDROEK 453
DB 187 ROMQEAIALLOGERRRGYRDGVAIGREI---FSQLPANNR---LRRREERWALENOR 238
OY 454 DDFIDDF-----SEEDYNDIVDGWKAKLVRTTEGQOW 486

```

DB 239 DECFADMYVHPQELDYN-----TETLRW 262

```

RESULT 12
UBIG_SALTY
ID UBIG_SALTY STANDARD; PRT; 209 AA.
AC P37431;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)
DE 3-DEMETHYLBIQUINONE-9-3-METHYLTRANSFERASE (EC 2.1.1.64)
DE (3,4-DIHYDROXY-5-HEXAPENTYLBENZOATE METHYLTRANSFERASE) (DHMB
DE METHYLTRANSFERASE) (FRAGMENT).
GN UBIG.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TF2;
RX MEDLINE=94253018; PubMed=8195103;
RA Jordan A., Gilbert I., Barbe J.;
RT "Cloning and sequencing of the genes from Salmonella typhimurium
RT encoding a new bacterial ribonucleotide reductase."
RL J. Bacteriol. 176;3420-3427(1994).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE +
3-DEMETHYLBIQUINONE-9 -> S-ADENOSYL-L-HOMOCYSTEINE +
BIQUINONE-9.
CC -1- PATHWAY: UBIQUINONE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE UBIG/COO3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X72948; CAA51451.1; -.
DR PIR; S32628; S32628.
DR SycGene; SG10405; ubiG.
KM Ubiquinone biosynthesis; Transferase; Methyltransferase.
FT NON TPR 1
SQ SEQUENCE 209 AA; 23061 MW; 3D07EB2B31D0D5E7 CRC64;

```

```

Query Match 4.9%; Score 127.5; DB 1; Length 209;
Best Local Similarity 26.4%; Pred. No. 0.017;
Matches 48; Conservative 34; Mismatches 63; Indels 37; Gaps 10;
OY 264 GYVS--TGLETTKEFVSKDLKPGOKVLDVCGGIGGDFYMAENVYD---VVGIDLSI 318
DB 13 GYIERRAGGLE-----GKRVLDVGC---GGG--ILASSMAREGATVAGLDMGF 55
OY 319 NMISFALERSIGLKCAVEFEVADCTKKDYP---ENSEVDIYSRDTILHIODKPALFRSFH 375
DB 56 EPILOYA--KLHALSFGIEVEYQETVEEHAHAKHAGQYDVYTCMELHEVDPQSVVRACA 113
OY 376 KWLKPGKVLISDYCKSA---GTPSAEFAVYIRORGYDLDHVKAYGKMKDKAGFEVYI 430
DB 114 QLVKPGGVFFSTLNRNGKSWLMAVGAELYILRMVPG--THDVK---KFKIPALLSWV 168
OY 431 AE 432
DB 169 DE 170

```

```

RESULT 13
Y829_SYNP3
ID Y829_SYNP3 STANDARD; PRT; 212 AA.

```

```

AC Q55423;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE PUTATIVE METHYLTRANSFERASE SLI0829 (EC 2.1.1.-).
GN SLI0829.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
ON NCBI_TaxID=1148;
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneke T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugiyama M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 648 to 928 of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64003; BAI0522.1; -.
DR InterPro: IPR000339; -.
DR Pfam: PF01209; Ublc_methyltran; 1.
KM Hypothetical protein; Transferase; Methyltransferase.
SQ SEQUENCE 212 AA; 23071 MW; BAD77E2E605A7001 CRC64;

Query Match 4.88; Score 126; DB 1; Length 212;
Best Local Similarity 27.7%; Pred. No. 0.022;
Matches 54; Conservative 32; Mismatches 77; Indels 32; Gaps 12;

QY 266 VSTGQLETTKEF-VSKIDLRGKQVLDVGGIGGGDFPMANVDEVWIDLSINMISFA 324
DB 25 LNVGGEERFQPLDLENLAISFGQVLDCCGGQATVYLADS-GATVGLDAS----- 76
QY 325 LENSIGLKCAVEFEVADCT-----KDYF--ENSGFYVSHDTLIHQDKPLFRSF--- 374
DB 77 -PKALG---RAKINVPQATVYQGLAEDLPREGSEFDLVHT-SVALH-EMTPAQLOSIISG 130
QY 375 -HKWLKFGKVLISDYKSAG---TPSAEF-AAVIRQGYDLHDVKA YGKMLKDGAVE 428
DB 131 VHRVLKPGGIFALVDLRHPSNMLFWPPLAIFMGLFETETAMOLINTD-LGSLIDQAGFTV 189
QY 429 VIAENRTDFOIYOLQ 443
DB 190 VRKHLVAGGSLOVLIQ 204

RESULT 14
TEHB_ECOLI
ID TEHB_ECOLI STANDARD; . PRT; 197 AA.
AC P25357; P76866;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE TELLURITE RESISTANCE PROTEIN TEHB.
GN TEHB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
ON NCBI_TaxID=562;
RX MEDLINE=91285417; PubMed=2060788;

```

```

RA Walter E.G., Weiner J.H., Taylor D.E.;
RT "Nucleotide sequence and overexpression of the tellurite-resistance
RT determinant from the IncHI1 plasmid pH1508a.";
RL Gene 101:1-7(1991).
RN 12]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.O.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN 13]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Osushima T., Saito N., Sampei G., Seki Y., Sivasubaram S.,
RA Tagami H., Takeeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horinouchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN 14]
RP GENE MAPPING.
RC STRAIN-K12;
RX MEDLINE=94222856; PubMed=8169225;
RA Taylor D.E., Hou Y., Turner R.J., Weiner J.H.;
RT "Location of a potassium tellurite resistance operon (teha tehb)
RT within the terminus of Escherichia coli K-12.";
RL J. Bacteriol. 176:2740-2742(1994).
CC -1- FUNCTION: RESPONSIBLE FOR POTASSIUM TELLURITE RESISTANCE WHEN
CC PRESENT IN HIGH COPY NUMBER, PROBABLY BY INCREASING THE REDUCTION
CC RATE OF TELLURITE TO METALLIC TELLURIUM WITHIN THE BACTERIUM.
CC OTHERWISE, PHENOTYPICALLY SILENT.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE PLASMID ENCODED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74072; AAA19564.1; -.
DR EMBL: AE000240; AAC74512.1; -.
DR EMBL: D90782; BAA15059.1; -.
DR EMBL: D90783; BAA15064.1; -.
DR PIR: JQ1018; JQ1018.
DR Ecogene: EG11884; tehb.
KM Antibioc resistance: Tellurium resistance.
SQ SEQUENCE 197 AA; 22531 MW; BF0D2D64F99C2E8 CRC64;

Query Match 4.88; Score 123.5; DB 1; Length 197;
Best Local Similarity 28.6%; Pred. No. 0.031;
Matches 59; Conservative 29; Mismatches 83; Indels 35; Gaps 13;

QY 254 ILRYRVFGGVSTGGLETTKEFVSKLD-LKPGQKVLVGGCIGGGDFYMAEN-YDVEY 311
DB 2 IIRDENFTDKYEITL---RTJSEVLEAVKVKPG-KTLDGGCCNGRNSITL LANGYDVA 57
QY 312 VGIDLSINMISFA-LER-SIGLKCAVEFEVADCTKKDYDENSEFDVYSHDTLIHQDK- 367
DB 58 -----WDKNAMSIANVERIKSIENLDLHTRVVDLNNLTF-DROYDFLSTVLMFLEAKT 112

```

QY 368 -PALFRSFHKVLKPGKVL-----SDYCKSAGTPSAEPAYIRQ--RGY----- 409
DB 113 IPGLIANMQRCTKPGYVLIYAAMDYADYPCYGVGFPAFKEGELRYYEGMERVKYNEDV 172
QY 410 -DLHDVYKYGKMLDAGFVEYIAENR 434
DB 173 GELHRTDANGNRIR-LRPATMLARKK 197

Job time: 123 sec

RESULT 15

Y095_HAEIN STANDARD; PRT: 251 AA.
AC 057060; 005007;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN HI0095.
CN HI0095.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE-95350630; PubMed-7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
Flint L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RT Science 269:496-512(1995).
RL -1 SIMILARITY: WEAK, TO YEAST ERG6 AND TO OTHER METHYLTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U32695; AAC21772.1; -.
DR TIGR; HI0095; -.
DR KW Hypothetical protein; Transferase; Methyltransferase.
SQ SEQUENCE 251 AA; 28050 MW; 75C3891A55D485B6 CRC64;

Query Match 4.7%; Score 123; DB 1; Length 251;
Best Local Similarity 25.6%; Pred. No. 0.047;
Matches 32; Conservative 30; Mismatches 57; Indels 6; Gaps 3;

QY 269 GGLETTKFEVSKIDLPKQGVLYDVGCGIGDFYMAENVYGVGIDLSINMISFA---L 325
DB 22 GKRATDWLILANGGFSQDKVLEACNMGTAIGLAKQFGCHIEGVDLDENALAKAQANI 81
QY 326 ERSIGLKCAVEFEVADCTKDKYDENSFDVIYSRD--TIIHIDDKPALFRSFHKVLKPGK 383
DB 82 EAN-GLOEKIHVQRANAKLPEDESFYIVNEMLTMLPVEAKKKAIAEYFRVLKPNGL 140
QY 384 VLISD 388
DB 141 LLTHD 145

Search completed: September 20, 2001, 16:21:26

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2001, 16:18:13 ; Search time 19.82 seconds
(without alignments)
3297.613 Million cell updates/sec

Title: US-09-525-885-2

Perfect score: 2598
Sequence: 1 MAASAMGVLOEREVEFKKYYI.....KIVRTGEQOQWGLFIKKM 494

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2598	100.0	494	10 Q9M571	Q9M571 spinacia ol
2	2105.5	81.0	491	10 Q9M571	Q9M571 spinacia ol
3	2093	80.6	498	10 Q9M571	Q9M571 spinacia ol
4	1287.5	49.6	374	10 Q9M571	Q9M571 spinacia ol
5	767.5	29.5	180	10 Q9M571	Q9M571 spinacia ol
6	498.5	19.2	437	5 Q22993	Q22993 caenorhabd
7	449.5	17.3	495	5 Q23552	Q23552 caenorhabd
8	223	8.6	565	2 Q9KJ20	Q9KJ20 actinopolys
9	207	8.0	363	10 Q41586	Q41586 tritlicum ae
10	207	8.0	363	10 Q41587	Q41587 tritlicum ae
11	204.5	7.9	317	2 Q55809	Q55809 synechocyst
12	199	7.7	278	2 Q9EY12	Q9EY12 streptomyc
13	196	7.5	346	10 Q82720	Q82720 nicotiana t
14	193.5	7.4	279	2 Q9KJ21	Q9KJ21 ectochloro
15	191	7.4	367	10 Q43445	Q43445 glycine max
16	189.5	7.3	275	2 Q9X508	Q9X508 streptomyc
17	189.5	7.3	283	2 Q9X509	Q9X509 streptomyc
18	189.5	7.3	387	2 Q9PNB2	Q9PNB2 campylobact
19	189	7.3	272	2 Q52570	Q52570 amycolatops

20	187.5	7.2	227	1 Q57965	Q57965 pyrococcus
21	187	7.2	336	10 Q9LM02	Q9LM02 arabidopsis
22	185	7.1	318	2 P74388	P74388 synechocyst
23	184	7.1	349	10 Q82434	Q82434 nicotiana t
24	184	7.1	438	2 Q9K558	Q9K558 streptomyc
25	183.5	7.1	462	2 Q9K557	Q9K557 streptomyc
26	182	7.0	344	10 Q49215	Q49215 zea mays (m
27	182	7.0	344	10 P93852	P93852 zea mays (m
28	179	6.9	278	2 Q9RNB8	Q9RNB8 microcystis
29	178	6.9	348	10 Q9ZSK1	Q9ZSK1 arabidopsis
30	177	6.8	348	10 Q9XIP9	Q9XIP9 arabidopsis
31	176.5	6.8	306	2 Q54095	Q54095 saccharopol
32	175.5	6.8	283	2 Q9S0N6	Q9S0N6 streptomyc
33	175.5	6.8	349	10 Q82426	Q82426 oryza sativ
34	174.5	6.7	227	1 Q9V268	Q9V268 pyrococcus
35	173	6.7	346	10 Q24328	Q24328 ricinus com
36	171	6.6	389	2 Q25171	Q25171 helicobacte
37	169.5	6.5	244	2 Q9EZC1	Q9EZC1 staphylococ
38	166	6.4	369	2 Q9ZK68	Q9ZK68 helicobacte
39	165.5	6.4	366	5 Q9YXP1	Q9YXP1 caenorhabd
40	163.5	6.3	254	2 Q9X036	Q9X036 thermotoga
41	162.5	6.3	581	10 Q64850	Q64850 arabidopsis
42	160	6.2	361	10 Q39227	Q39227 arabidopsis
43	160	6.2	361	10 Q9LN25	Q9LN25 arabidopsis
44	159.5	6.1	286	2 P94922	P94922 mycobacteri
45	159.5	6.1	391	1 Q52025	Q52025 halobacteri

ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	494 AA.
ID	Q9M571			
AC	Q9M571			
DT	01-OCT-2000 (T-EMBLrel. 15, Created)			
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)			
DE	PHOSPHOETHANOLAMINE N-METHYLTRANSFERASE (EC 2.1.1.103).			
GN	PEAMT.			
OS	Spinacia oleracea (Spinach).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;			
OC	Caryophyllales; Chenopodiaceae; Spinacia.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Salinized LEAF;			
RX	MEDLINE=20261526; PubMed=10799484;			
RA	Nuccio M.L., Ziemak M.J., Henry S.A., Weretillyn E.A., Hanson A.D.;			
RT	"Phosphoethanolamine N-methyltransferase from spinach: cDNA cloning by complementation in schizosaccharomyces pombe and characterization of the recombinant enzyme";			
RL	J. Biol. Chem. 275:14095-14101(2000).			
DR	EMBL; AF237633; AAF61950.1; -			
DR	InterPro; IPR000051; -			
DR	InterPro; IPR000780; -			
KW	Transferase; Methyltransferase.			
SQ	SEQUENCE 494 AA; 56361 MW; 7F2537C8EAB8413B CRC64;			
Query Match	100.0%; Score 2598; DB 10; Length 494;			
Best Local Similarity	100.0%; Pred No. 1.4e-168;			
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 MAASAMGVLOEREVEFKKYYIHSVDLTFEAMMLDSQASDDKVERPEVLSMPPECKSV 60			
DB	1 MAASAMGVLOEREVEFKKYYIHSVDLTFEAMMLDSQASDDKVERPEVLSMPPECKSV 60			
QY	61 LEAGAGIGRTGTELAERASQVIALDFIESYIKKNSINGHYKKNKFCACVYTSPLNISP 120			
DB	61 LEAGAGIGRTGTELAERASQVIALDFIESYIKKNSINGHYKKNKFCACVYTSPLNISP 120			

QY	121	NSVDIIIFSWMILMYLSDSEVERBLVERBMLKWLKPGCIIFERESCFHOSGDKRKSNTFHYR	180
QY <td>121</td> <td>NSVDIIIFSWMILMYLSDSEVERBLVERBMLKWLKPGCIIFERESCFHOSGDKRKSNTFHYR</td> <td>180</td>	121	NSVDIIIFSWMILMYLSDSEVERBLVERBMLKWLKPGCIIFERESCFHOSGDKRKSNTFHYR	180
Db <td>121</td> <td>NSVDIIIFSWMILMYLSDSEVERBLVERBMLKWLKPGCIIFERESCFHOSGDKRKSNTFHYR</td> <td>180</td>	121	NSVDIIIFSWMILMYLSDSEVERBLVERBMLKWLKPGCIIFERESCFHOSGDKRKSNTFHYR	180
QY <td>181</td> <td>EBRFYTKIFKECHMODDSGNSYELSLIGCKICGAYVKSNNQOISWLMQKVSDEDDKGF</td> <td>240</td>	181	EBRFYTKIFKECHMODDSGNSYELSLIGCKICGAYVKSNNQOISWLMQKVSDEDDKGF	240
Db <td>181</td> <td>EBRFYTKIFKECHMODDSGNSYELSLIGCKICGAYVKSNNQOISWLMQKVSDEDDKGF</td> <td>240</td>	181	EBRFYTKIFKECHMODDSGNSYELSLIGCKICGAYVKSNNQOISWLMQKVSDEDDKGF	240
QY <td>241</td> <td>QRFLLDSOYKFNLSILRYERFPGPYVSTGGLETTKEFVSKLDLPGQKVLVDVCGIGGGD</td> <td>300</td>	241	QRFLLDSOYKFNLSILRYERFPGPYVSTGGLETTKEFVSKLDLPGQKVLVDVCGIGGGD	300
Db <td>241</td> <td>QRFLLDSOYKFNLSILRYERFPGPYVSTGGLETTKEFVSKLDLPGQKVLVDVCGIGGGD</td> <td>300</td>	241	QRFLLDSOYKFNLSILRYERFPGPYVSTGGLETTKEFVSKLDLPGQKVLVDVCGIGGGD	300
QY <td>301</td> <td>PYMAENYDVEYVIGDILSTIMISFALERSIGLCAVEFEVADCTKDYPENSDYIYSRPT</td> <td>360</td>	301	PYMAENYDVEYVIGDILSTIMISFALERSIGLCAVEFEVADCTKDYPENSDYIYSRPT	360
Db <td>301</td> <td>PYMAENYDVEYVIGDILSTIMISFALERSIGLCAVEFEVADCTKDYPENSDYIYSRPT</td> <td>360</td>	301	PYMAENYDVEYVIGDILSTIMISFALERSIGLCAVEFEVADCTKDYPENSDYIYSRPT	360
QY <td>361</td> <td>ILIHIDKRALPFRSFKHKWLKPGGKVLISDYCKSAGPSAFEAAYIROGYDLHDVKAYGKM</td> <td>420</td>	361	ILIHIDKRALPFRSFKHKWLKPGGKVLISDYCKSAGPSAFEAAYIROGYDLHDVKAYGKM	420
Db <td>361</td> <td>ILIHIDKRALPFRSFKHKWLKPGGKVLISDYCKSAGPSAFEAAYIROGYDLHDVKAYGKM</td> <td>420</td>	361	ILIHIDKRALPFRSFKHKWLKPGGKVLISDYCKSAGPSAFEAAYIROGYDLHDVKAYGKM	420
QY <td>421</td> <td>LKDAFVEVYIAENRTDOEIQVLQKELDALDQEKDDDFIDFSESDYNDIYDGKAKLVRTT</td> <td>480</td>	421	LKDAFVEVYIAENRTDOEIQVLQKELDALDQEKDDDFIDFSESDYNDIYDGKAKLVRTT	480
Db <td>421</td> <td>LKDAFVEVYIAENRTDOEIQVLQKELDALDQEKDDDFIDFSESDYNDIYDGKAKLVRTT</td> <td>480</td>	421	LKDAFVEVYIAENRTDOEIQVLQKELDALDQEKDDDFIDFSESDYNDIYDGKAKLVRTT	480
QY <td>481</td> <td>EGEQOMGLEFIARKM 494</td> <td></td>	481	EGEQOMGLEFIARKM 494	
Db <td>481</td> <td>EGEQOMGLEFIARKM 494</td> <td></td>	481	EGEQOMGLEFIARKM 494	
RESULT	2		
Q9FR44	Q9FR44	PRELIMINARY: PRT: 491 AA.	
AC	Q9FR44:		
DT	01-MAR-2001 (Tremblrel. 16, Created)		
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)		
DT	01-MAR-2001 (Tremblrel. 16, Last annotation update)		
DE	SAV: PHOSPHO-ETHANOLAMINE N-METHYLTRANSFERASE.		
GN	NMT1.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
OC	Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_taxid=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. LANDSBERG ERECTA; TISSUE-STAGE TWO LEAVES;		
RX	MEDLINE=20567827; PubMed=11115895;		
RA	Biologenie C.P., McGraw P.;		
RT	"The Isolation and Characterization in Yeast of a Gene for Arabidopsis		
RL	s-adenosylmethionine:phospho-ethanolamine N-methyltransferase.";		
RL	Plant Physiol. 124:1800-1813(2000).		
DR	EMBL; AF197940; AAC41121.1.;		
DR	Transferase; Methyltransferase.		
QO	SEQUENCE 491 AA: 56102 MW; DA4404EBBD3FA8D5 CRC64;		

Query Match	81.0%;	Score 2105.5;	DB 10;	Length 491;
Best Local Similarity	78.9%;	Pred. No. 4e-135;		
Matches 389;	Conservative 55;	Mismatches 46;	Indels 3;	Gaps 1.

QY	1	MAASAMGYLOEEREVAKKRWIEHSVDTLVEAMMLDSQASDLDERVEREYSMTLPPEYGSKV	60
Db	1	MAAS---YEERNDIKNNWIEHSADLTVEAMMLDSRASDLDEREYVSLTLPPEYGSKV	57
QY	61	LELGAGIRFTGELAEKASQVIALDFIESVYIKNBSINCHYKNVAKFMCADVTSPSLINSP	120
Db	58	LELGAGIRFTGELAKQAGELTALDFIDIVAKIKNSINCHYKNVAKFMCADVTSPDLKITD	117
QY	121	NSVDLIEFSNMLIMTSDSEVERLVERMLKWLKPGYIFPFECSGFQSGHKKKSNPTHR	180
Db	118	GSLLDIFSNMLIMTSDSEVELLARRMGMKIVGVIIFPFECSGFQSGHKKKSNPTHR	177

[illegible]

ID	09LVH3	PRELIMINARY;	PRT;	498 AA.
AC	Q9LVH3;			
AD	01-OCT-2000 (TReMBLrel. 15, Created)			
DT	01-OCT-2000 (TReMBLrel. 15, last sequence update)			
DT	01-MAR-2001 (TReMBLrel. 16, last annotation update)			
DE	METHYL TRANSFERASE-LIKE PROTEIN.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosid II;			
OC	Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=COLUMBIA.			
RA	Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;			
RL	Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=COLUMBIA.			
RA	MEDLINE=20277480; PubMed=10819329;			
FX	Nakamura Y.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence			
RT	features of the regions of 4,504,864 bp covered by sixty p1 and TAC			
RT	clones.";			
RL	DNA Res. 7:131-135(2000).			
DR	EMBL; AB019230; BAB02720.1; -.			
DR	InterPro; IPR000051; -.			
DR	InterPro; IPR001601; -.			
KW	Transferase.			
SO	SEQUENCE 498 AA; 56877 MW; E57693B915ECF0D8 CRC64;			

```

Query Match Similarity      80.0%; Score 2093; DB 10; Length 498;
Best Local Similarity      78.0%; Pred. No. 2, 9e-134;
Matches 390; Conservative 54; Mismatches 46; Indels 10; Gaps 2

QY      1  MAASAMGVLQREVERKYYIEHSVDLTVEAMMLDSQASDLDRVEREVLSTLPPYEGKSV 60
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db      1  MAAS---YEEEDIDQKNYIEHSADLTVEAMMLDSASDLDEEREPVSLTPPYEGKSV 57
QY      61  LELGAGIGFPGELAEKASOVATLDFIESVIKKNSINGHYKNVRCMDVYSPSLNISP 120
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db      58  LELGAGIGFPGELAKQAGELLALDFIDVNIKKNSINGHYKNVRCMDVYSPDLKTTD 117
QY      121 NSVDIIFSWMLLMYLSDEVERLVRLMKLWMLPGGYIFPESCFPHOSGDHKRRKSNETHYR 180
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```


Db	118	GSUJLFSNMWLTAMWLSQKEVELLAERWVGNKVGVIFFRESCFHQSGDKRSKNPTPHR	177
Qy	181	EPFRYTKI-----FKBCHMODSGNSYSLSLGICKCICAYVSKKNOQISWLMQKD	233
Db	178	EPREFSKIDGETMKVFECCQTRDAAGNSFELSMIGCKICAIYAVKKNKNOQICWIMQKS	237
Qy	234	SEDDKGFQRFELDSQYKFNSLTRERERFGCGVYSTGGLFTTKFEVSKLDLCPQKVLVVG	293
Db	238	SENRGFRQFRLDNQYKSSGLTRERERFGCGVYSTGGLFTTKFEVEMNLKPOQKVLVVG	297
Qy	294	CGIGGGFFYMAENADVEVAVGIDISINMISFALERSIKLCKAVEFVADCTKKYPPNSFD	353
Db	298	CGIGGGFFYMAEKRDVHAYVIGDLSVNMISFALERAISLSCSVEFVADCTTKHYPPNSFD	357
Qy	354	VIYSRDTFLIHLQDKPALFRSEFHWKLKDGKGVILSDYCKSAGTSAEFAAYIRQGYDLHD	413
Db	358	VIYSRDTFLIHLQDKPALFRFPEFKWLKDGKGVILSDYCRSPKPTSAESEYIKRQGYDLHD	417
Qy	414	VKATGKMLKADGFEVIAENRTDFIOYLQKELDLEQEKDPTIDFSEEDYDIYDGMK	473
Db	418	VQAGOMGLKADGFEVDVIAEDRTDFQMOYLRKELDRVEKEKEKTSIFSKEEDYDIVGMK	477
Qy	474	AKLVRTTEGEOMGLFTAKK	493
Db	478	SKLERCASDEQKGLFTANK	497
RESULT	4		
ID	Q9LP63	PRELIMINARY;	PRT; 374 AA.
AC	Q9LP63;		
DT	01-OCT-2000 (TReMBLrel. 15, Created)		
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)		
DT	01-MAR-2001 (TReMBLrel. 16, Last annotation update)		
DE	TM15.23.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
OC	Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_Taxid=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,		
RA	Shinn P., Altafi H., Bei Q., Chin C., Chlou J., Choi E., Conn L.,		
RA	Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,		
RA	Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M.,		
RA	Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayeri A.,		
RA	Toriiumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,		
RA	Ecker J.R.;		
RT	'Genomic sequence for Arabidopsis thaliana BAC T1N15 from chromosome		
RT	1.';		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Ecker J.R.;		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Ecker J.R.;		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Chauk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,		
RA	Khan S., Kim C., Altafi H., Bei B., Chin C., Chlou J., Choi E.,		
RA	Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,		
RA	Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,		
RA	Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,		
RA	Thayeri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,		
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: ACO20889; AAF7905.1; -.		
DR	InterPro: IPR0000051; -.		
DR	InterPro: IPR001601; -.		

SEQ	SEQUENCE	374 AA:	42656 MW:	318D85E03110C047	CNC64:
Query Match		49.6%:	Score 1287.5:	DB 10:	Length 374:
Best Local Similarity		66.4%:	Pred. No. 9.6e-80:		
Matches	243:	Conservative	36:	Mismatches	28:
				Indels	59:
				Gaps	2:
QY	187	KTFKCHMDDSGNSYEYSLGCKICIGAVYVSKKNONQISWIMQVDSDDKGFQRIIDS	246		
DB	8	GVFOECOTRDASGNSFELSMWGCKICIGAVYVSKKNONQISWIMQVDSDDKGFQRIIDS	67		
QY	247	SOYKRNSTLRVRYRGPVSGSL-----ETKRE	276		
DB	68	VOYKSSGILRERVRGEGVSTGFGNSILLSSYHTLYCLSVIMFLSTYETKRE	127		
QY	277	FVSKSLDKPGQVLDVGGCIGGDPYMAENYDVEVYGDLSINMISFALENSIGLKCAVE	336		
DB	128	FVAKKMDLKPQGVLDVGGCIGGDPYMAENYDVEVYGDLSINMISFALERIGLKCAVE	187		
QY	337	FEVADCTKKDVPENSFVYIYSHDTLLIHODKPALEFRSFHMKLPGKVLISDYCKSAGTP	396		
DB	188	FEVADCTKKDVPENSFVYIYSHDTLLIHODKPALEFRSFHMKLPGKVLITDYCRSAETP	247		
QY	397	SAEFAAYIRQKGYDLHDYKAYGKMLKADGFEVYVIAENRTDQFIQYLQKELDALDEKEDF	456		
DB	248	SPEFAEYIKQKGYDLHDYQAYGMKLDGFDVYIAEDRDQFVQVLRRELEKVEKEKEEF	307		
QY	457	IDDFSE-----EDYNDIYDGKKAKLVRTTEGEOQMG	487		
DB	308	ISDFSEVKTFLTTNSFEONTCLLCSHIDHKILLLOEDYNDIYDGWSAKLERTASGEQKMG	367		
QY	488	LEFIARK 493			
DB	368	LEFIADK 373			
RESULT	5				
ID	Q9LP64	PRELIMINARY:	PRT:	180 AA.	
AC	Q9LP64:				
DT	01-OCT-2000 (TREMBlrel. 15, Created)				
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)				
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)				
DT	TIM15.22.				
OC	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eurariopsis: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:				
OC	Magnoliophyta: eudicotyledons, core eudicots; Rosidae; eurosids II;				
OC	Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,				
RA	Shinn P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,				
RA	Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,				
RA	Leuz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,				
RA	Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayerl A.,				
RA	Toriunli M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,				
RA	Ecker J.R.;				
RT	"Genomic sequence for Arabidopsis thaliana BAC T1N15 from chromosome				
RT	I.,"				
RL	Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Ecker J.R.;				
RL	Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Ecker J.R.;				
RL	Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Chen R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,				
RA	Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,				

RA Conn L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharasy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southlake A.,
 RA Thayer A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.,
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC020889; AAF79704.1; -
 DR InterPro: IPR000051; -
 DR InterPro: IPR000780; -
 DR InterPro: IPR001601; -
 SQ SEQUENCE 180 AA; 20472 MW; AAOFEFD2DD27E556 CRC64;

Query Match 29.5%; Score 767.5; DB 10; Length 180;
 Best Local Similarity 75.4%; Pred. No. 7e-45;
 Matches 144; Conservative 19; Mismatches 17; Indels 11; Gaps 1;

QY 20 IEHSVDLTFEAMMDSQASDLKVERPEVLSMLPPEGKSYLELGAGIGFTGLAEKAS 79
 Db 1 MEHSDDLTFEAMMDSKASDLKEREPEVLSLIPPEGKSYLELGAGIGFTGLAEKAG 60
 QY 80 QVTLADTFESYIKKESINGHYKVKFMCADVTSPSLNISPSVDITFSNMLVLSDEE 139
 Db 61 EVTLADTFESYIKKESINGHYKVKFMCADVTSPDLKIKGSDILFSSNMLVLSDEE 120
 QY 140 VERLVERMLKWLKPGYIFFRESCFHSGDHKRSNPTHYREPRFYTKIFKCHMODSG 199
 Db 121 VELMAERHMGVKGPGYIFFRESCFHSGDSKRSNPTHYREPRFYTKV----- 169

QY 200 NSYSLSLGCK 210
 Db 170 QSEDLTFELFK 180

RESULT 6
 022993 PRELIMINARY; PRT; 437 AA.

AC Q22993; 01-NOV-1996 (TREMUR1.01, Created)
 DT 01-NOV-1996 (TREMUR1.01, Last sequence update)
 DT 01-JUN-2000 (TREMUR1.14, Last annotation update)
 DE SIMILAR TO METHYLTRANSFERASES.
 GN F54D11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Pauley A., Gattung S.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64834; AAB04824.1; -
 DR InterPro: IPR000051; -
 DR InterPro: IPR000339; -
 DR InterPro: IPR001601; -
 DR Pfam: PF01209; Ublc_methyltran; 1.
 KW Transferase; Methyltransferase.
 SQ SEQUENCE 437 AA; 49769 MW; 6675E262F627D08B CRC64;

Query Match 19.2%; Score 498.5; DB 5; Length 437;
 Best Local Similarity 32.6%; Pred. No. 4.e-26;
 Matches 127; Conservative 70; Mismatches 156; Indels 37; Gaps 10;

QY 121 NSVYIFSNMLVLSDEEVLVERMLKMLK-----GGYIFFRESCFHSGDHK 173
 Db 64 NADVLIFSNMLVLSDEEVLVERMLKMLK-----LTDPLKNTATAGIYIRED-LKDSKRV 116
 QY 174 SNPHYREPRFYTKIFKCHMODSGNSYELSGCKI--GAYVSKKNQNSIWM-- 229
 Db 117 ARLDY-----FVFRITDSDGNNTGLDLYTVQVESHNVY-----QNFDFIFV 163
 QY 230 -QKV---DSEDDKGFQRLDSSQYKFNLSILRYEAVFPGYSTGLTFEYK-LDK 284
 Db 164 RKVFAPPTDATTFRDLDTQYNTGIDAYEMFVNFSPGVDENLTIIRFGDFK 223
 QY 285 PGQKVLVGGGIGGDEYMAENVYDVEVVDLSTNMISFALERSIGLK-CAVEFEVADCT 343
 Db 224 PGQTMIDIGVIGGAGQVADDFGVHVGIDLSNMLALALERHREKDSVKSITDAL 283
 QY 344 KKDYPENSFDVYISRDITLHIDQKPLFRSFHKLKPGKVLISDYCSACTPAEPAAY 403
 Db 284 VYQFEDNSFDVYFSRDCIOHIDPEKLFRIYKALKPGKVLIMYKGGVGEQSDKFKTY 343
 QY 404 IROGVDLHDVKAQKMLKMGFEVIAENRTDFOIYLOKELALDQEKDFIDDSSE 463
 Db 344 VAQRATFLKMLKETADIRANKTGFVNVOTENMTPEFKLEERGHLEONEAEFSKFTOR 403
 QY 464 DYNDIVDGMKAKLVRTTEGEQMGFLIAKK 493
 Db 404 ERDSLISGWDKLGYLEKDNHNMNFFLAOK 433

RESULT 7

023552 PRELIMINARY; PRT; 495 AA.

AC Q23552; 01-NOV-1996 (TREMUR1.01, Created)
 DT 01-NOV-1996 (TREMUR1.01, Last sequence update)
 DT 01-JUN-2000 (TREMUR1.14, Last annotation update)
 DE CODED FOR BY C. ELEGANS CDNA YK92B11.3.
 GN ZK622.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RP [2]


```

OY 280 KDLDPGKVLVDYGGIGCGDFYMAENYDVGJDISINNIS--FLNERSIGLKAVEF 337
      ::::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|
Db 114 QLELPGKVLVDYGGIGCGDFPLREIARESTYSTVGLNNNDYITGKALNSVGLGATCDF 1737
OY 338 EVADCTKKDYPENSPFDVYISDITLHIDOKPALFRSPHKWLKPGKVLIDSYC-KSAGTP 396
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 174 VAKDEMPKPFSDNFDVAYIAEATCHADPVGCKEYIRVLKPGQCFAYVEWCTIDHYDP 233
OY 397 SAEFAAYR---QKRYDLHDYKANGKML---KDAQFEVLT-----AENR----- 433
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 234 NNATKRKRKDEIEIENGSLPDIRSTROCLOAVKADGAF-EVIDMKDLAEPSLPWLYPLDPS 292
OY 435 -----TQGFIOVLQKELDALEQEKDEIDDFSPREDYNDYDGKAKL 476
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 293 RESLSPRLTGVGRIITRNMKVVL-EYVGAPLPGSOSRVSSFLERKAEGLEVGGKKEI 348

RESULT 10
Q41587
AC Q41587 PRELIMINARY; PRT; 363 AA.
Dt 01-NOV-1996 (TREMBLrel. 01, Created)
Dt 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
Dt 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE DELTA-24-STEROL METHYLTRANSFERASE.
GN TA-MT.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;
OC Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Subramaniam K., Ueng P.P.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; 060755; AAB49338.1; -
DR Mendel; 10874; Triae; 1138; 10874.
DR InterPro; IPR000051; -
KW Transferase; Methyltransferase.
SQ SEQUENCE 363 AA; 41636 MW; 7F20995861a62c75 CRC64;

```

Query Match	Similarity	8.0%	Score	207	DB	10	Length	363
Best Local	Similarity	28.6%	Pred. No.	2,2e+06				
Matches	68	Conservative	38	Mismatches	88	Indels	44	Gaps

QY	280	KIDLPKGRKVLDDVGGCGIGGGDFYMAENYDVEYVGLDLSINMTS	-FALERSIGLKCAVEF	337
Db	114	OLELPGKMLVDVGGCGIGGIPGLREIRAFSSTSVGLNNNDYQITRGKALNRSVGLGATCDF		173
QY	338	EVADCTKDDPENSDVIVYSRPTIIHIDQKPLKPLRSFHKMLPGGKVLISDVC	-KSAGNP	396
Db	174	VKADMKMPFSNDTDAVYATATCATCHADPDCVQKREIVYVLKPGQCFAVEYWCITDHYDP		233
QY	397	SAEPAAYLR--QRCYDLHDVAKVYKML--KDAGEFVEYI-----AENR-----		434
Db	234	NNATHKRLKDEIEIENGLEPDIRSTQCCQAVKADAGF-EVIMPKDLALSDSPLPWYLPLDPS		292
QY	435	-----TDQFIQVLOKELDALFQEKDFIDFDSSESDYNDIVYGKAKTL		476
Db	293	RFSLSFRITTVGRITTRNMVKVL--EYVGLAPESQSRVSSFLERAKAEGIVGGRKEI		348

RESULT	11
Q55809	
AC	Q55809 PRELIMINARY: PRT: 317 AA.
DT	01-NOV-1996 (TREMBlrel. 01, Created)
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE	DELTA1(24)-STEROL C-METHYLTRANSFERASE.
GN	ERG6 OR SLR0089.
OS	Synechocystis sp. (strain PCC 6803).

XX Bacteria;Cyanobacteria; Chroococcales; Synechocystis.
 OC NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugita M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64k to 92k of the genome.";
 RL DNA Res. 2:153-166(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hoshouchi T., Matsuno A., Muraki A., Nakazaki N., Nairu K., Okumura S.,
 RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D64004; BAI0562.1; -.
 DR InterPro; IPR000051; -.
 DR InterPro; IPR001601; -.
 KW Transferase; Methyltransferase.
 SO SEQUENCE 317 AA; 35571 MW; 98DD01C081B89F72 CRC64;

	Query Match	7.98;	Score 204.5;	DB 2;	Length 317;
	Best Local Similarity	26.38;	Pred. No. 2,7e+06;		
	Matches 61;	Conservative 48;	Mismatches 90;	Indels 33;	Gaps
Qy	215	YKSKKNQNOISWLMQKVDSEDDKGFQRFILDSOQKFNISILRRYRPGYVSTGLET	274		
Db	35	YEKIKNFEDDSGGLMEDVWGE-----HMHGYYGPHGTVRIDR-----RQAOIDLI	80		
Qy	275	KEF-----VSKIDLPGQKVLIDVGGIGGGDFPYMAENVDEYVGLDLSINMISFAL	--RS 328		
Db	81	KELAMAVPQNSAP-RKILIDLGCGIGGSSLYTLAQOHAELWAGSLSVQVERGERARA	139		
Qy	329	IGLCAVEFEVADCTKKDYPENSDVYISRDITLHIDKRALPFSFHKWLPGQKVLISD	388		
Db	140	LGLSTQCFQVYANMLDLPFADSPSDWWSLSGSHMKNKQFLQELARVYLKPGGRLLAT	199		
Qy	389	YC-----KSACTPSAEPAATYRKQGYDLH-----DYKANGKMLKDXGFEV	429		
Db	200	WCHRPIDPGNGPLTADERRHL-QAITYDVYCLPYVVSPLDYEAIRREGCFGEI	250		
RESULT	12				
QyEY12	ID	QyEY12	PRELIMINARY;	PRT;	278 AA.
AC	QyEY12;				
DT	01-MAR-2001 (TREMBlrel. 16, Created)				
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)				
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)				
DE	SNOM.				
GN	SNOM.				
OS	Streptomyces nogalater.				
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
OC	Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.				
OX	NCBI_TaxID=38314;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC27451;				
RA	Torkkeli S., Kunnari T., Palmu K., Mantasala P., Hakala J.,				
RA	Ylihanko K.;				
RT	"The entire nogalamyacin biosynthetic gene cluster of Streptomyces				
RT	nogalater; Characterization of a 20 kb DNA region and generation of				
RT	hybrid structures.";				
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.				

DR EMBL; AF323753; AAG42853.1; -.
SQ SEQUENCE 278 AA; 30466 MW; 9804BC2024BD85C1 CRC64;

Query Match 7.7%; Score 199; DB 2; Length 278;
Best Local Similarity 26.6%; Pred. No. 5.2e-06;
Matches 62; Conservative 48; Mismatches 101; Indels 22; Gaps 8.

```

QY 274 TKRVSXLDLPGKVLVDGCGGGGGFYMAENVDVVGIDISINNISPA---LEESIG 330
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 TEMMTKEIENGPGORVLIDIGCGICAPAVOLARATGAEVGVTITISPEOVLATAHAERE-G 110
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 331 LKCAVEEVDADCTCK-DYPENSEPDVIXSRDTILHIDOKPALFRSEFHKMLPGKG-----V 384
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 VAEVTFRCADASAHELDPADSPFDAVMFPESIFLPLDRALTRAAEVLPGGRLALTDV 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 385 LISDYCSAGTSPSAEFPAAYIRQRTGDLHDYKAYKMLKDGFVEY-----IAENRTDQFIQ 440
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 LHNETSPPAQOESLDEHAHYTPLVGEPMR-LSDPYPLLRLQARLVPVECRDISELTVGRGLE 229
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 441 VLQKELALBEOKDFIDDFSEEDYNDIVOGWKALVRTTEGEQOMGLFIATKK 493
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 CMHHTL-----ENRERYERGGSE-----LVDFSTAVLLLAVEGGYIVTPAAR 275
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT	13	
082720		
ID	082720	PRELIMINARY; PRT; 346 AA.
AC	082720;	
DT	01-NOV-1998 (TREMBLrel. 08, Created)	
DT	01-NOV-1998 (TREMBLrel. 08, last sequence update)	
DT	01-JUN-2000 (TREMBLrel. 14, last annotation update)	
DE	5-ADENOSYL-METHIONINE-STEROL-C- METHYLTRANSFERASE.	
OS	Nicotiana tabacum (Common tobacco)	
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	
OC	Magnoliophyta; eudicotyledons; Core eudicots; Asteridae; euasterids I.	
OC	Solanales; Solanaceae; Nicotiana.	
OX	NCBI_TaxID=1097;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	SPRAIN-XANTHI SH6.	
RX	MEDLINE=98417432; PubMed=9746350;	
RA	Bouvier-Nave P., Husselstein T., Benveniste P.;	
RT	"Two families of sterol methyltransferases are involved in the first	
RT	Eur. J. Biochem. 256:88-96(1998).	
DR	EMBL: U81312; AAC34951.1; "	
DR	Mendel: 33290; Nicita:1138;33290.	
DR	InterPro: IPR000051; "	
KM	Transferase; Methyltransferase.	
SQ	SEQUENCE 346 AA; 38871 MW; E7A3F9534A9D94AA CRC64;	

[illegible]

AC G9KJ21;
DT 01-OCT-2000 (Tremblrel. 15, created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE SARCOSINE-DIMETHYLGLYCINE METHYLTRANSFERASE.
OS Ectothiorhodospira halochloris.
OC Bacteria; Proteobacteria; gamma subdivision; Ectothiorhodospiraceae;
OC Halorhodospira
OX NCBI_TaxID=1052;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20357342; PubMed=10896953;
RA NYssola A., Kervouo J., Kaukinen P., von Weymar N., Reinkainen T.
RT "Extreme halophilic synthetize betaine from glycine by methylation."
RJ J. Biol. Chem. 275:22196-22201(2000).
DR EMBL: AF216282; AAF87203.1; -.
DR InterPro: IPR000051; -.
DR Interpro: IPR001601; -.
KW Transferase; Methyltransferase.
SQ SEQUENCE 279 AA; 32227 MW; 579337A69156605 CRC64;

[illegible]

RESULT	15			
043445				
ID	043445	PRELIMINARY:	PRT:	367 AA.
AC	043445			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-JUN-2000	(TREMBLrel. 14, Last annotation update)		
DE	S-ADENOSYL-L-METHIONINE: DELTA24-STEROI-C- METHYLTRANSFERASE.			
OS	Glycine max (Soybean).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I;			
OC	Fabales; Fabaceae; Papilionoideae; Glycine.			
OX	NCBI_Taxid=38477;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=CV. WILLIAMS 82; TISSUE=ETIOLATED HYPOCOTYLS;			
RC	MEDLINE=96199190; Pubmed=8621604;			
RX	Shi J., Gonzales R.A., Bhattacharyya M.K.;			
RA	"Identification and Characterization of an S-adenosyl-L-methionine			
RT	delta 24-sterol-C-methyltransferase cDNA from soybean."			
RL	J. Biol. Chem. 271:9384-9389(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. WILLIAMS 82; TISSUE=ETIOLATED HYPOCOTYLS;			
RA	Clouse J.A.;			
RL	Submitted (DEC-1995) to the EMBI/Genbank/DBJ databases.			

